

На правах рукописи



ЛЕВИН Евгений Калманович

**РАЗРАБОТКА СРЕДСТВ ИССЛЕДОВАНИЯ И ПОВЫШЕНИЯ
ПОМЕХОУСТОЙЧИВОСТИ СИСТЕМ АВТОМАТИЧЕСКОГО
РАСПОЗНАВАНИЯ ГОЛОСОВЫХ КОМАНД В ТЕЛЕФОНИИ**

Специальность 05.12.13 – Системы, сети и устройства телекоммуникаций

АВТОРЕФЕРАТ

диссертации на соискание ученой степени
доктора технических наук

Владимир 2014

Работа выполнена на кафедре радиотехники и радиосистем
Федерального государственного бюджетного образовательного учреждения высшего
профессионального образования «Владимирский государственный университет
имени Александра Григорьевича и Николая Григорьевича Столетовых»

Научный консультант

Никитин Олег Рафаилович, доктор технических наук, профессор
заведующий кафедрой радиотехники и радиосистем Федерального государственного
бюджетного образовательного учреждения высшего профессионального об-
разования «Владимирский государственный университет имени Александра Гри-
горьевича и Николая Григорьевича Столетовых»

Официальные оппоненты:

Семенов Андрей Борисович, доктор технических наук
директор по развитию ООО РдМ Дистрибьюшн, г. Москва.

Фархадов Маис Паша Оглы, доктор технических наук, старший научный
сотрудник заведующий лабораторией Федерального государственного бюджет-
ного учреждения науки «Институт проблем управления им. В. А. Трапезникова
Российской академии наук», г. Москва.

Шаврин Сергей Сергеевич, доктор технических наук, доцент
профессор Московского технического университета связи и информатики,
г. Москва.

Ведущая организация

ОАО «Владимирское конструкторское бюро радиосвязи», г. Владимир.

Защита диссертации состоится «11» ноября 2014 г. в 14.00 часов на заседа-
нии диссертационного совета Д 212.025.04 при Владимирском государственном
университете имени Александра Григорьевича и Николая Григорьевича Столето-
вых по адресу: г. Владимир, пр-т Строителей, 3/7, ауд. 301-3.

С диссертацией можно ознакомиться в научной библиотеке Владимирского
государственного университета имени Александра Григорьевича и Николая Гри-
горьевича Столетовых.

Автореферат разослан «30» июля 2014 г.

Отзывы на автореферат, заверенные печатью, просим направлять по адресу:
600000, г. Владимир, ул. Горького, 87, ВлГУ, ФРЭМТ.

Ученый секретарь диссертационного совета
доктор технических наук, профессор



А. Г. Самойлов

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы

Привлекательность речевого «общения» абонента телефонной сети с автоматическими информационными системами, успехи научных исследований автоматического распознавания речи, а также развитие вычислительной техники обусловили в настоящее время развитие систем голосового самообслуживания (СГС). СГС состоит из трех основных составляющих: интерактивного автоответчика, системы автоматического распознавания голосовых команд (САРГК) и устройства синтеза речи. Основными областями использования СГС являются контактные центры и справочно-информационные службы.

Использование СГС позволяет освободить операторов контактных центров от часто повторяющихся вопросов и тем самым снизить их психологическую нагрузку, обеспечить круглосуточную работу справочной службы и быстрое обновление информации для клиентов.

Необходимость развития СГС обусловило процесс перехода систем автоматического распознавания голосовых команд (САРГК) из научных лабораторий в сферу практической деятельности человека.

Особенностью САРГК является работа в условиях сильной изменчивости произнесения команд. Изменчивость команд успешно учитывается при формировании акустических моделей отдельных звуков и команд. В качестве моделей обычно используются модели скрытых марковских процессов (МСМП). В отсутствие помех достигнута достаточно малая относительная частота ошибок распознавания приемлемая для использования САРГК на практике.

Однако наличие акустических, электрических и радиопомех обуславливает большое количество ошибок распознавания, что препятствует широкому применению САРГК. Воздействие помех приводит либо к ложному срабатыванию системы, либо к ошибке в распознавании команды.

Большой вклад в решение проблемы повышения достоверности автоматического распознавания речи внесли следующие ученые: Бондарко Л.В., Винцюк Т.К., Галунов В.И., Гейлс М., Грей А., Маркел Дж.Д., Потапова Р.К., Прохоров Ю.Н., Рабинер Л.Р., Сапожков М.А., Фант Г., Фланаган Дж., Хуанг К., Шафер Р.В., Янг Б. Работы данных исследователей и их последователей позволили значительно снизить частоту ошибок распознавания. При проведении научных исследований получены приемлемые для практики результаты.

Однако внедрение САРГК в практическую деятельность человека наталкивается на большие трудности. Из-за разнообразия помех сложно учесть их влияние в конкретных условиях эксплуатации системы при формировании акустических моделей команд на стадии разработки САРГК. Приходится проводить дополнительную статистическую обработку большого количества звукозаписей с учетом наличия помех при настройке САРГК для работы в конкретных условиях эксплуатации. Проводится большое количество экспериментов для создания звукозаписей и их дальнейшей обработки. Поэтому создание помехоустойчивых САРГК требует больших затрат.

Эксперименты проводятся, во-первых, с целью определения параметров алгоритмов обработки речевого сигнала (РС), обеспечивающих требуемую

достоверность распознавания команд, а во-вторых, для оценки параметров достоверности. После каждого эксперимента определяются относительные частоты ложного срабатывания системы (False Alarm Rate - FAR) и ложного пропуска команды (False Reject Rate - FRR), которые являются оценками соответствующих вероятностей и характеризуют достоверность распознавания. Чем выше требования к достоверности распознавания, тем больше объем тестовых выборок звукозаписей и выше трудозатраты.

Особо следует отметить проблему определения FAR. В связи с многообразием помех затруднительно создавать достаточно большую тестовую базу для каждого конкретного случая использования САРГК. Обычно создается некоторая большая тестовая звуковая база, на которой тестируются различные системы. Полученное значение FAR оказывается «привязанным» к данной тестовой базе. Так как помехи в конкретных условиях эксплуатации системы, в общем случае, отличаются от тех, которые отражены в тестовой базе, то значение FAR, определенное при тестировании системы, может не соответствовать значению FAR в условиях эксплуатации системы. На этапе внедрения системы приходится проводить сбор дополнительных данных о помехах и особенностях эксплуатации системы для более точной ее настройки.

С расширением сферы использования САРГК растет количество пользователей систем и увеличивается разнообразие помех, воздействующих на системы. Расширение круга пользователей требует повышения достоверности распознавания команд, что наряду с ростом разнообразия помех приводит к увеличению объема экспериментов при разработке САРГК. Соответствующее увеличение трудозатрат еще можно реализовать в рамках научных исследований, но при инженерном проектировании систем такие затраты часто становятся «неподъемными».

Становится **актуальной** проблема создания средств исследования помехоустойчивости САРГК, обеспечивающих снижение объема экспериментов при выборе параметров алгоритмов обработки РС и оценке достоверности распознавания команд. Снижение объема экспериментов целесообразно обеспечить разработкой соответствующих методов оценки достоверности распознавания, а также средств оценки степени подавления помех без привлечения мощных вычислительных ресурсов.

Существующие средства разработки, в основном, направлены на создание акустических моделей голосовых команд, реализацию алгоритмов сопоставления параметров речевых сигналов с акустическими моделями и определение параметров речевых сигналов (РС). Средствам исследования алгоритмов подавления помех уделено мало внимания. Особенно это касается алгоритмов подавления помех на стадии предварительной обработки речевых сигналов до процедуры определения параметров сигналов, используемых при распознавании.

Объектом исследования являются системы автоматического распознавания голосовых команд в телефонии.

Предметом исследования являются средства исследования и повышения помехоустойчивости систем автоматического распознавания голосовых команд в телефонии.

Целью диссертационной работы является создание средств исследования помехоустойчивости САРГК, которые обеспечивают сокращение объема экспериментов при разработке помехоустойчивых САРГК для нужд телефонии, а также создание средств повышения помехоустойчивости САРГК.

Для достижения поставленной цели в диссертационной работе необходимо **решить следующие задачи.**

Провести анализ существующих методов обеспечения и исследования помехоустойчивости САРГК.

Разработать методы исследования помехоустойчивости САРГК, обеспечивающие снижение объема экспериментов при их проектировании.

Проанализировать возможности и разработать алгоритмы подавления помех на стадии предварительной обработки сигналов.

Разработать программные и аппаратные средства для исследования алгоритмов подавления помех.

Научная новизна характеризуется следующими результатами диссертационной работы.

- Предложен критерий достоверности распознавания команд, определяемый при тестировании САРГК выборкой звукозаписей.

- Предложен метод сравнения результатов распознавания команд, полученных при разных сеансах тестирования САРГК в процессе их проектирования, по достоверности распознавания.

- Предложен метод оценки вероятности ложного срабатывания системы в зависимости от степени соответствия посторонних произнесений акустической модели ключевого слова команды и получены результаты экспериментального исследования метода.

- Получены выражения, отражающие зависимость вероятности ошибки сравнения результатов распознавания по достоверности распознавания от объема тестовой выборки звукозаписей и параметров достоверности сравниваемых систем.

- Получены выражения, отражающие зависимость степени подавления квазипериодических помех методом неадаптивной компенсации от соотношения уровней помехи, шума и речевого сигнала.

- Получены выражения, отражающие зависимость степени подавления широкополосных радиопомех методом адаптивной компенсации от погрешности реализации требуемых значений задержек в адаптивных трансверсальных фильтрах.

- Предложен алгоритм адаптивной компенсации радиопомех с подавлением влияния сигнала на процесс адаптации путем использования обратной связи по решению.

Теоретическая значимость работы заключается в следующем.

- Проведен анализ метрики сопоставления произнесения с акустическими моделями помехи, «своей» и «чужой» команд;

- оценена вероятность принятия ошибочного решения при сравнении САРГК по достоверности распознавания;

- проанализировано влияние аддитивного шума на результат оценки периода помехи при компенсации квазипериодических помех;

– получено выражение оптимального вектора управляющих коэффициентов для процессора Хоуэлса-Эпплбаума, используемого для адаптивной компенсации широкополосных радиопомех с отражениями;

– исследовано влияние обратной связи по решению, используемой для уменьшения влияния сигнала на работу цепи адаптации, на степень подавления помехи с помощью адаптивной компенсации.

Практическая значимость работы заключается в следующем.

– Использование предложенного критерия сравнения результатов распознавания команд, полученных при разных сеансах тестирования САРГК, по достоверности распознавания позволяет уменьшить объем тестовой выборки звукозаписей;

– с целью снижения объема экспериментов, разработана методика выявления помехи и голосовой команды, которым соответствует нижняя граница достоверности распознавания;

– разработана методика оценки вероятности ложного срабатывания САРГК на произнесения слов, не являющихся ключевыми словами команд, которая позволяет снизить объем экспериментов при настройке системы на конкретные условия эксплуатации;

– разработаны программные средства для получения оценки вероятности ложного срабатывания, слабо зависящей от состава тестовой выборки звукозаписей;

– разработан набор функций системы Matlab для исследования алгоритмов компенсации помех на основе имитационного моделирования;

– разработаны программно-аппаратные средства, обеспечивающие сбор данных о появлении ошибок распознавания команд в процессе эксплуатации САРГК на телефонной линии;

– разработаны структурные схемы комплекса аппаратуры для испытаний устройств компенсации радиопомех.

Внедрение

1. Результаты диссертационной работы используются в ООО «Телеком-Сервис» для оценки качества услуг, предоставляемых в телефонном контакт-центре.

2. Предложенный метод сравнения САРГК, использован в ООО «Центр Речевых Технологий» при проектировании системы VoiceDigger автоматического поиска ключевых слов в записях телефонных разговоров.

3. Предложенный метод оценки вероятности ложного срабатывания САРГК на произнесения посторонних слов был использован в ООО «НПП Акустика» при разработке системы «Барышня» голосовой навигации.

4. Разработанные программные средства, используются на кафедре радиотехники и радиосистем Владимирского государственного университета при проведении лабораторных работ в рамках учебных дисциплин «Компьютерная телефония» и «Цифровая обработка сигналов изображения и звука».

Методы исследования

При проведении исследований использовались методы математической статистики, матричного анализа, математический аппарат теории случайных функций, адаптивной фильтрации, а также имитационное моделирование устройств обработки сигналов и экспериментальное исследование систем распознавания голосовых команд.

На защиту выносятся

1. Метод сравнения результатов распознавания команд по достоверности распознавания, который использует предложенный критерий достоверности, основанный на определении выборочного среднего и выборочной дисперсии.
2. Метод оценки вероятности ложного срабатывания с учетом степени несоответствия произнесения постороннего слова акустической модели команды.
3. Результаты теоретического анализа и имитационного моделирования компенсации квазипериодических помех на основе оценки периода помехи.
4. Результаты теоретического анализа и имитационного моделирования устройств адаптивной компенсации широкополосных радиопомех.
5. Результаты теоретического анализа устройств адаптивной компенсации радиопомех с использованием обратной связи по решению для подавления влияния сигнала на процесс адаптации.

Достоверность результатов исследований обусловлена использованием соответствующего математического аппарата, имитационного моделирования и экспериментальной проверкой.

Апробация результатов работы

Основные положения диссертации докладывались на 2-10 МНТК «Перспективные технологии в средствах передачи информации», (1997-2013г.), Владимир-Суздаль; на 2,4-8,10 МНТК «Физика и радиоэлектроника в медицине и экологии», (1996-2012г.), Владимир-Суздаль; на 4-8 межрегиональных НТК «Обработка сигналов в системах двусторонней телефонной связи» (1995-1998), Москва-Пушкинские горы, МТУСИ, на 10, 13 межрегиональных НТК «Обработка сигналов в системах телефонной связи и вещания», (2000, 2004 г.), Москва-Пушкинские горы; МТУСИ, на международной конференции SPECOM'2003. – Moscow State Linguistic University, Moscow, Russia, 2003, на МНПК «Фундаментальные проблемы радиоэлектронного приборостроения INTERMATIC-2004», (2004 г.), Москва; на международной конференции SPECOM'2005, – University of Patras, Patras, Greece, 2005.

Публикации

По теме диссертации опубликовано 70 работ, в том числе одна монография, 26 статей, из них 13 статей в изданиях, рекомендованных ВАК, 4 авторских свидетельства на изобретения, 39 тезисов докладов в трудах международных и российских конференций.

Личное участие

Основные теоретические результаты были получены лично автором в период с 1974 по 2014 г. Разработка программных и аппаратных средств, а также экспериментальные исследования проводились коллективом сотрудников при активном участии автора в ходе выполнения научно-исследовательских работ,

проводимых в качестве исполнителя и научного руководителя на кафедре радиотехники и радиосистем Владимирского государственного университета.

Структура и объем работы

Диссертация состоит из Введения, шести разделов, Заключения, списка литературы, имеющего 207 наименований отечественных и зарубежных источников, в том числе 70 работ автора, и Приложений. Общий объем диссертации составляет 257 страниц, в том числе 201 страницы основного текста, 33 страницы списка литературы, 50 рисунков, 16 таблиц и 22 страницы приложений.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во Введении обоснована актуальность работы, сформулированы цели и задачи исследований, дана характеристика научной новизны и практической значимости результатов работы, приведены данные о структуре диссертации.

В первом разделе диссертации дана краткая характеристика основных этапов обработки речевых сигналов (РС) при их распознавании. Сначала с помощью детектора голосовой активности абонента проводится удаление пауз из сигнала. Затем сигнал разбивается на короткие сегменты, соответствующие интервалам квазистационарности РС, и для каждого сегмента определяется набор параметров, которые характеризуют огибающую кратковременного спектра сигнала. В качестве параметров РС наиболее часто используются мелочастотные кепстральные коэффициенты (Mel-Frequency Cepstral Coefficients - MFCC), которые дополняются своими первыми и вторыми производными.

Полученная последовательность наборов (векторов) параметров сопоставляется с акустическими моделями команд, которые состоят из акустических моделей звуков. Чтобы учесть значительную изменчивость произнесения звуков голосовой команды (ГК), в качестве модели каждого звука используется модель скрытого марковского процесса (МСМП).

Для учета возможности произнесения различных команд и наличия пауз используется сеть акустических моделей. В процессе сопоставления некоторого произнесения с сетью моделей определяется та последовательность моделей, которая соответствует произнесению с наибольшей вероятностью. Из-за изменчивости произнесения ГК и наличия помех возникают ошибки распознавания.

Источником помех, во-первых, является сам диктор, который непроизвольно издает посторонние звуки и произносит посторонние слова, не являющиеся ГК. Во-вторых, помехи создает акустическое окружение диктора. В-третьих, источниками помех являются средства связи.

Наличие пауз в РС обусловило широкое распространение метода спектрального вычитания для подавления помех. Оценивается спектральная плотность мощности помехи на интервале паузы с последующим вычитанием ее из спектральной плотности РС, пораженного аддитивной помехой. Недостатком метода является наличие помехи типа «музыкальный» шум.

В Европейском стандарте ETSI ES 202 212 V1.1.2, посвященном созданию САРГК для телефонии, подавление помех осуществляется с помощью фильтра

Винера, который минимизирует искажения сигнала, возникающие при его фильтрации.

Различные методы подавления помех используются для обработки РС с целью их последующего прослушивания. Здесь, в частности, используется компенсация помех. При компенсации параметры РС не искажаются как в случае использования фильтров, но отношение сигнал-шум снижается, так как к широкополосному шуму, который присутствует в сигнале, добавляется шум, присутствующий в паузе сигнала.

Подавление помех, которые находятся вне интервала произнесения командного (ключевого) слова, осуществляется и на стадии сопоставления параметров РС с акустическими моделями. Для помех создаются акустические модели. Для подавления произнесений посторонних слов создают подсеть акустических моделей отдельных звуков, которая включается в общую сеть моделей.

Воздействие помех снижает достоверность распознавания команд – увеличиваются вероятности ложного пропуска команды и ложного срабатывания системы на помеху. Помехоустойчивость САРГК оценивают по степени изменения указанных вероятностей. Широко используемыми оценками вероятностей являются относительные частоты ложного пропуска команды (False Reject Rate - FRR), и ложного срабатывания системы (False Alarm Rate - FAR). Указанные оценки определяются в ходе тестирования САРГК объемными выборками соответствующих звукозаписей. Чтобы обеспечить малую погрешность оценок, необходимо при росте достоверности распознавания увеличивать объем тестовой звуковой базы, что приводит к росту затрат на разработку системы.

Величина относительной частоты ложного срабатывания зависит от состава тестовой выборки звукозаписей - звуковой базы. Из-за отличий звукозаписей тестовой базы от произнесений, распознаваемых при эксплуатации системы, требуется проведение дополнительного тестирования с помощью специально создаваемой звуковой базы.

Проведенный анализ методов подавления помех показывает, что на предварительной стадии обработки сигналов наименьшее изменение параметров РС обуславливает использование методов компенсации. Целесообразно исследовать его возможности для различных случаев воздействия помех.

В процессе проектирования САРГК приходится неоднократно проводить ее тестирование для оценки вероятности ошибки распознавания команды. Для обеспечения высокой точности сравнения результатов нескольких процедур тестирования необходимо использовать объемные тестовые выборки звукозаписей. Причем с уменьшением вероятности ошибки распознавания приходится увеличивать объемы тестовых выборок.

Целесообразно разработать методы оценки вероятностей ошибки распознавания команд и ложной тревоги, которые обеспечивают снижение объема экспериментальных исследований при проектировании САРГК.

Во втором разделе поведен анализ метрик сопоставления параметров РС с акустическими моделями, входящими в сеть моделей. Рассматривается сеть, в которой модели произнесений ключевых слов образованы последовательным

соединением акустических моделей звуков – трифонов. Под трифоном, соответствующим какому-либо звуку, понимается совокупность трех соседних звуков, где данный звук является центральным. Модели, соответствующие посторонним произнесениям, состоят из моделей монофонов. Монофон – это звук без учета окружения соседними звуками.

Результат сопоставления некоторого произнесения с моделью, входящей в состав сети, можно отобразить следующей последовательностью символов: 0 0 0 0 1 1 1 1 1 1 0 0 0 0. Каждый символ соответствует результату сопоставления произнесенного звука с некоторой моделью звука. Если звук соответствует модели, то имеет место символ единицы. В противном случае присутствует символ нуля.

При сопоставлении звука с монофонной моделью возможно лишь два варианта результата сопоставления: соответствует звук модели или не соответствует. Обозначим через $e_{m,j}$ метрику сопоставления для первого результата, а через $d_{m,j}$ – метрику второго варианта. Здесь индексы m, j обозначают соответственно монофонный тип модели и j -ый номер звука в произнесении.

Если звук полностью соответствует трифонной модели с учетом контекстов, то имеет место метрика $e_{tRL,j}$. Индекс RL обозначает, что слева и справа звук окружают звуки, которые соответствуют его трифонной модели. Если звук полностью не соответствует модели, то имеет место метрика $d_{t,j}$. Если модель и звук соответствуют друг другу, но отличаются по одному из контекстов (левому или правому) или по обоим контекстам, то таким ситуациям соответствуют метрики: $e_{tL,j}$ (или $e_{tR,j}$) и $e_{t,j}$. Метрика сопоставления всего произнесения с моделью ГК равна сумме метрик сопоставления звуков с соответствующими моделями.

Зная метрики сопоставления звуков произнесения с моделями звуков и зная количество метрик, можно определить метрики сопоставления произнесения с любой моделью из сети моделей. Решение о результате распознавания произнесения принимается путем определения модели, обеспечивающей максимальную метрику сопоставления с произнесением.

Рассмотрим разность метрик сопоставления ГК со «своей» трифонной и монофонной моделями, считая, что каждый звук ГК сопоставляется с соответствующей моделью.

$$\Delta M_{t-m,ii} = M(\mathbf{X}_i | \Lambda_{t,i}) - M(\mathbf{X}_i | \Lambda_{m,i}) = e_{m,1} + \sum_{j=2}^{W-1} e_{t,j} + e_{m,W} - \sum_{j=1}^W e_{m,j} = \sum_{j=2}^{W-2} (e_{t,j} - e_{m,j}) = \sum_{j=2}^{W-2} \Delta e_{t-m,j}.$$

Здесь $\Delta e_{t-m,j}$ - разность метрик сопоставления матрицы параметров j -го звука с трифонной и монофонной моделями; W – число моделей звуков в модели команды; \mathbf{X}_i – матрица параметров произнесения i -ой ГК; $\Lambda_{t,i}$ – набор моделей звуков данной команды. В процессе тестирования САГК распознаванию подвергается множество произнесений ГК, поэтому величина $\Delta M_{t-m,ii}$ является случайной величиной равной сумме случайных величин $\Delta e_{t-m,j}$, которые зависят от множества случайных вариаций длительности и тембра произнесения j -го звука. Можно считать, что ее распределение $p(\Delta M_{t-m,ii})$ плотности вероятностей близко нормальному закону. Если $\Delta M_{t-m,ii} > 0$, то произнесенная ГК распознается верно.

Если же $\Delta M_{t-m,ii} < 0$, то ГК трактуется как помеха - имеет место ошибка распознавания.

Математическое ожидание $\mu_{t-m,ii}$ и дисперсия $\sigma_{t-m,ii}^2$ величины $\Delta M_{t-m,ii}$ равны сумме математических ожиданий и дисперсий случайных величин $\Delta e_{t-m,j}$.

$$\mu_{t-m,ii} = \sum_{j=2}^{W-2} \mu_{e,t-m,ij}, \quad \sigma_{t-m,ii}^2 = \sum_{j=2}^{W-2} \sigma_{e,t-m,ij}^2, \quad (1)$$

где $\mu_{e,t-m,ij}$, $\sigma_{e,t-m,ij}^2$, математическое ожидание и дисперсия величины $\Delta e_{t-m,j}$. Учтем, что $\mu_{e,t-m,ij} > 0$, так как модель трифона, в среднем, обеспечивает большую метрику сопоставления со звуком по сравнению с моделью монофона. Следовательно, вероятность того, что ГК будет распознана как постороннее произнесение

$$P_{FR,i} = \int_{-\infty}^0 p(\Delta M_{t-m,ii}) d(\Delta M_{t-m,ii}) = 0,5 - \Phi(\alpha), \quad (2)$$

где $\Phi(\alpha) = \frac{1}{\sqrt{2\pi}} \int_0^\alpha e^{-z^2/2} dz$ - функция Лапласа, $\alpha = \mu_{t-m,ii}/\sigma_{t-m,ii}$. С учетом (1), видно, что с ростом числа W звуков в ГК растет величина α , и уменьшается вероятность $P_{FR,i}$ ошибки распознавания.

Если метрика сопоставления произнесения постороннего слова с совокупностью монофонных моделей его звуков больше метрики сопоставления с моделью k -го ключевого слова, то имеет место правильное распознавание – система относит это произнесение к числу помех. В противном случае имеет место ошибка ложного срабатывания системы. Рассмотрим разность соответствующих метрик.

$$\Delta M_{m-t,ik} = M(\mathbf{X}_i | \Lambda_{m,i}) - M(\mathbf{X}_i | \Lambda_{t,k})$$

Указанная разность образована суммой следующих типов разностей: $e_m - d_m$, $e_m - d_t$, $e_m - e_{tR(L)}$, $e_m - e_{tRL}$. Все эти разности являются случайными величинами. Математические ожидания разностей первого и второго типов имеют положительные значения, четвертого – отрицательное. Математическое ожидание разности третьего типа может быть как отрицательным, так и положительным.

Результат распознавания зависит, в основном, от соотношения сумм математических ожиданий разностей второго и четвертого типов. Если рассматривать наиболее вероятную ситуацию, когда сумма математических ожиданий всех разностей, определяющих метрику $\Delta M_{m-t,ik}$, больше нуля, то вероятность ложного срабатывания системы определяется выражением (2), где величина α является отношением указанной суммы математических ожиданий к суммарному среднеквадратическому отклонению.

Так как величина метрики сопоставления зависит от степени соответствия произнесения звука модели, то математическое ожидание величины ΔM_{im-GKj} состоит из суммы отрицательных и положительных слагаемых. К положительным слагаемым относятся математические ожидания, соответствующие следующим типам разностей метрик: $(e_m - d_t)$, $(e_m - d_m)$. Отрицательные слагаемые: $(e_m - e_{tRL})$, $(e_m - e_{tL})$, $(e_m - e_{tR})$, $(e_m - e_t)$.

Чем больше метрик типа d и меньше метрик типа e , тем больше отношение α в выражении (2) и, следовательно, меньше вероятность ложного срабатывания.

Соотношение чисел метрик различных типов характеризует вероятность ложного срабатывания.

Проведенный анализ метрики сопоставления произнесения с моделями команд и помех показывает, что вероятности ложного срабатывания системы и распознавания произнесенной данной ГК в виде другой команды во многом определяются числом звуков произнесения, которые соответствуют трифонным моделям звуков, входящим в состав какой-либо модели ГК.

На основе проведенного анализа метрики предлагается следующий порядок оценки вероятности ложного срабатывания. Для заданной ГК при тестировании определяются те фрагменты слитно произносимых слов звуковой базы, которые вызвали ложное срабатывание САГК. Для каждого фрагмента определяется структура произнесений, которые могут вызвать ложное срабатывание. В структуре указываются те звуки произнесения (транскрипционные знаки), которые совпали со знаками транскрипции ГК. Кроме того, помечаются те знаки произнесения, которые отличаются от соответствующих знаков транскрипции ГК.

Далее определяются в транскрипции звуковой базы все фрагменты, которые соответствуют данной структуре. Определяется величина FAR как отношение числа фактических ложных срабатываний к числу фрагментов, которые соответствуют данной структуре. Для каждой структуры определяется своя величина FAR. Аналогично определяется FAR для каждой ГК из заданного их перечня.

Для получения обобщенной оценки целесообразно определять величину FAR и для кода метрики, который определяется количеством типов метрики сопоставления звука с акустической моделью. В этом случае набор значений FAR для соответствующих кодов метрик характеризует устойчивость САГК к воздействию расширенного ряда посторонних произнесений относительно данной ГК. Аналогичную обобщенную оценку можно использовать и для всего набора заданных команд.

Введем числовые параметры, характеризующие степень отличия в произнесениях ключевого и неключевого слов:

w – число транскрипционных знаков ключевого слова;

a – число одинаковых транскрипционных знаков ключевого и неключевого слов;

e – число одинаковых транскрипционных знаков ключевого и неключевого слов, имеющих слева и справа одинаковое окружение;

s – число одинаковых знаков ключевого и неключевого слов, имеющих одинаковое окружение только с одной стороны;

d – число одинаковых знаков ключевого и неключевого слов, имеющих разное окружение с каждой стороны;

i – число ложных тревог (срабатываний);

g – число фрагментов звуковой базы, «созвучных» ключевому слову и соответствующих данной структуре помехи.

Совокупность параметров $\{w, e, s, d\}$ определяет степень «созвучности» ключевого и неключевого слов, поэтому от нее зависит обобщенная оценка вероятности ложной тревоги для группы ключевых слов с числом w транскрипционных знаков.

Для оценки вероятности ложного срабатывания на новом звуковом материале проводится анализ транскрипции звукового материала, на котором планируется использовать систему распознавания, и оценивается вероятность появления различных структур помехи. С использованием оценок, полученных при тестировании, определяется оценка вероятности ложного срабатывания на данном звуковом материале.

Предложенный метод оценки вероятности ложной тревоги проверен экспериментально. Определялась интервальная оценка вероятности ложного срабатывания для каждого кода структуры постороннего произнесения по результатам тестирования САРГК с заданным списком ключевых слов двумя звуковыми базами («База 3», «База 4»). Количество произнесенных слов - 19068 для первой базы и 16776 для второй базы.

На Рисунке 1 представлено графическое отображение интервальных обобщенных оценок вероятности ложной тревоги для заданного перечня ключевых слов в зависимости от набора параметров $\{w, e, s, d\}$. Условные обозначения на графике: \circ – База 3, $*$ – База 4. По вертикальной оси отложены значения десятичных логарифмов границ интервалов, а по горизонтальной - номера наборов.

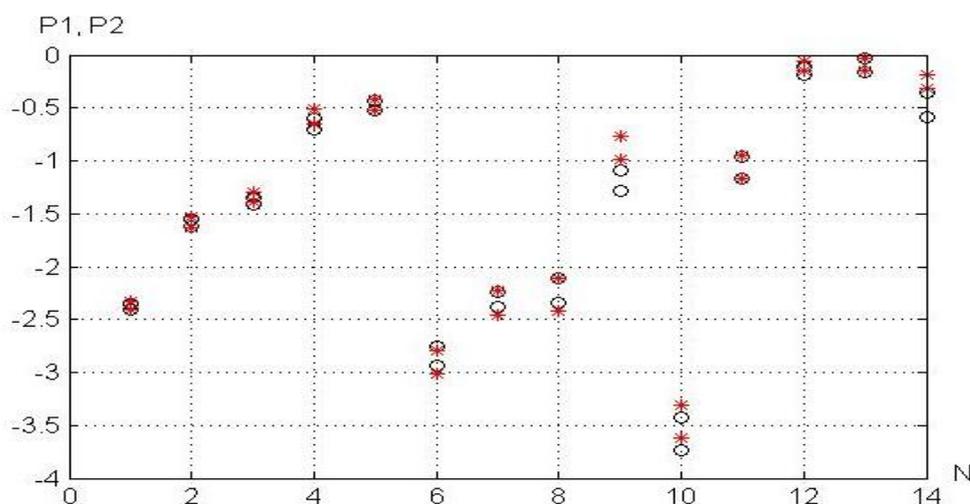


Рисунок 1. Доверительные интервалы оценки вероятности ложного срабатывания для списка ключевых слов

Так как База 3 больше Базы 4 на 2292 слова, то интервалы оценок вероятности для нее уже. Рост объема базы приводит к увеличению числа оценок с узкими интервалами (соответствует количеству ложных срабатываний для каждой структуры помехи большому 10) примерно в два раза. Перекрывание интервальных оценок свидетельствует о слабой зависимости оценок от состава звуковой базы. Следовательно, их можно использовать для расчета вероятности ложного срабатывания на звуковом материале, соответствующем условиям эксплуатации системы.

При разработке САРГК приходится неоднократно сравнивать результаты распознавания, относящиеся к разным сессиям тестирования системы, по достоверности распознавания. В качестве критерия достоверности распознавания обычно используется относительная частота ошибки распознавания, которая

является случайной величиной. Поэтому возможно принятие ошибочного решения по результатам сравнения. Чем выше достоверность распознавания, тем больше должна быть тестовая база, чтобы вероятность ошибки сравнения сохранялась на заданном уровне.

Предлагается для снижения объема тестовой выборки звукозаписей использовать критерий достоверности распознавания, который определяется следующим образом. При оценке вероятности ошибки распознавания i -ой команды в процессе тестирования САРГК, определяется метрика M_{ii} сопоставления текущего произнесения команды со «своей» моделью, а также - наибольшая метрика M_{ij} сопоставления данного произнесения с j -ой «чужой» моделью. Определяется разность метрик $\Delta M_i = M_{ii} - M_{ij}$.

Если $\Delta M_i > 0$, то результат распознавания считается правильным. В противном случае имеет место ошибка распознавания. Величина разности ΔM_i определяется изменчивостью произнесения i -ой команды. Изменчивость, в свою очередь, определяется множеством факторов: временной и тембральной изменчивостью произнесения, отличиями голосов разных дикторов, воздействием помех. Поэтому распределение вероятностей случайной величины ΔM_i можно считать нормальным. Данное утверждение подтверждено проверкой результатов экспериментов по критерию «хи-квадрат».

Если определить оценки $\hat{\mu}$ математического ожидания и среднеквадратического отклонения $\hat{\sigma}$ для данного распределения, то, используя отношение $\hat{\alpha} = \hat{\mu}/\hat{\sigma}$ при сравнении результатов распознавания, можно сократить требуемое число произнесений этой команды в тестовой звуковой базе по сравнению с известным методом. При этом вероятность принятия ошибочного решения по результатам сравнения систем не уменьшается.

Исходя из данного выше определения случайной разности ΔM_i , вероятность $P_{\text{ош.ГК}}$ ошибки распознавания команды

$$P_{\text{ош.ГК}} = \int_{-\infty}^0 p(\Delta M_i) d(\Delta M_i) = 0,5 - \Phi(\alpha),$$

где $\Phi(\alpha) = \frac{1}{\sqrt{2\pi}} \int_0^\alpha e^{-z^2/2} dz$ - функция Лапласа; $\alpha = \mu/\sigma$; μ , σ - математическое ожидание и среднеквадратическое отклонение нормального распределения, соответственно; $p(\Delta M_i)$ - плотность нормального распределения вероятностей случайной величины ΔM_i . Из приведенной формулы следует, что с увеличением α снижается вероятность ошибки распознавания.

Покажем, что использование предложенного критерия по сравнению с известным критерием снижает требуемый объем тестовой базы. Пусть результаты распознавания САРГК2 соответствуют меньшей вероятности ошибки распознавания по сравнению с результатами САРГК1, тогда $\Delta\alpha = \alpha_2 - \alpha_1 > 0$. На практике можно использовать лишь статистические оценки $\hat{\alpha} = \hat{\mu}/\hat{\sigma}$. Если $\Delta\hat{\alpha} = \hat{\alpha}_2 - \hat{\alpha}_1 < 0$, то возникает ошибка сравнения САРГК.

Проанализируем, насколько достоверным является сравнение САРГК по относительной частоте \bar{p} возникновения ошибки распознавания. Рассмотрим случайную величину $\Delta\bar{p} = \bar{p}_1 - \bar{p}_2$, где \bar{p}_1 , \bar{p}_2 - относительные частоты возникновения ошибки распознавания для первой САРГК1 и второй САРГК2,

соответственно. Для определенности считаем, что вторая САРГК2 обеспечивает более достоверное распознавание ГК, то есть вероятность ошибки для второй САРГК меньше, чем для первой: $p_2 < p_1$. Если при сравнении САРГК случайная величина $\Delta\bar{p}$ окажется положительной, то будет принято верное решение: САРГК2 обеспечивает большую достоверность по сравнению с САРГК1. В противном случае принимается ошибочное решение по результатам сравнения САРГК. Определим вероятность того, что $\Delta\bar{p}$ примет отрицательное значение при соблюдении условия: $p_2 < p_1$. Вероятность принятия ошибочного решения при сравнении САРГК по относительной частоте возникновения ошибки

$$p_{p.ош} = P(-1 < \Delta\bar{p} < 0) = \Phi\left(\frac{0-\Delta p}{\sigma_{\Delta p}}\right) - \Phi\left(\frac{-1-\Delta p}{\sigma_{\Delta p}}\right) = 0.5 - \Phi\left(\frac{\Delta p}{\sigma_{\Delta p}}\right).$$

Если задана вероятность принятия ошибочного решения, то, пользуясь таблицей значений функции Лапласа, из данного выражения можно определить отношение $\frac{\Delta p}{\sigma_{\Delta p}}$ и, задаваясь ожидаемыми значениями вероятностей p_1 и p_2 , можно определить требуемый объем тестовой выборки

$$n_p = \left(\frac{\Delta p}{\sigma_{\Delta p}}\right)^2 \frac{p_1(1-p_1) + p_2(1-p_2)}{\Delta p^2}$$

Вероятность принятия ошибочного решения при сравнении САРГК и при использовании предлагаемого критерия

$$P_{ош.сравн} = 0,5 - \Phi(\alpha_\beta), \alpha_\beta \approx \sqrt{n} \frac{\alpha_2 - \alpha_1}{\sqrt{2 + 0,5(\alpha_1^2 + \alpha_2^2)}}.$$

При заданной вероятности $P_{ош.сравн}$ ошибочного решения и ожидаемых значениях вероятностей p_1 , p_2 ошибок распознавания для САРГК1 и САРГК2 можно определить требуемый объем тестовой выборки звукозаписей

$$n_\alpha \approx \alpha_\beta^2 \frac{2 + 0,5(\alpha_1^2 + \alpha_2^2)}{(\alpha_2 - \alpha_1)^2}.$$

Полученные значения были проверены с помощью моделирования в среде Matlab. Построены графики (Рисунок 2) зависимостей вероятностей ошибочного решения при использовании известного и предложенного методов (на графике – P_{err}), (n – объем тестирующей выборки) для различных значений вероятностей P_1 , P_2 ошибок распознавания голосовых команд для САРГК1 и САРГК2, соответственно.

Здесь знаком «*» помечены данные, полученные при использовании параметра α , знак «o» относится к данным, полученным на основе определения относительной частоты ошибки распознавания. Графики являются достаточно гладкими, что говорит о малой величине погрешностей статистических оценок вероятностей ошибок сравнения, полученных при моделировании.

Анализ графиков показывает, что с увеличением объема тестирующей выборки уменьшается вероятность ошибки сравнения САРГК и растет выигрыш (уменьшение объема выборки) при применении предложенного метода сравнения САРГК. Выигрыш также увеличивается при уменьшении степени отличия сравниваемых САРГК по вероятности ошибки распознавания команд.

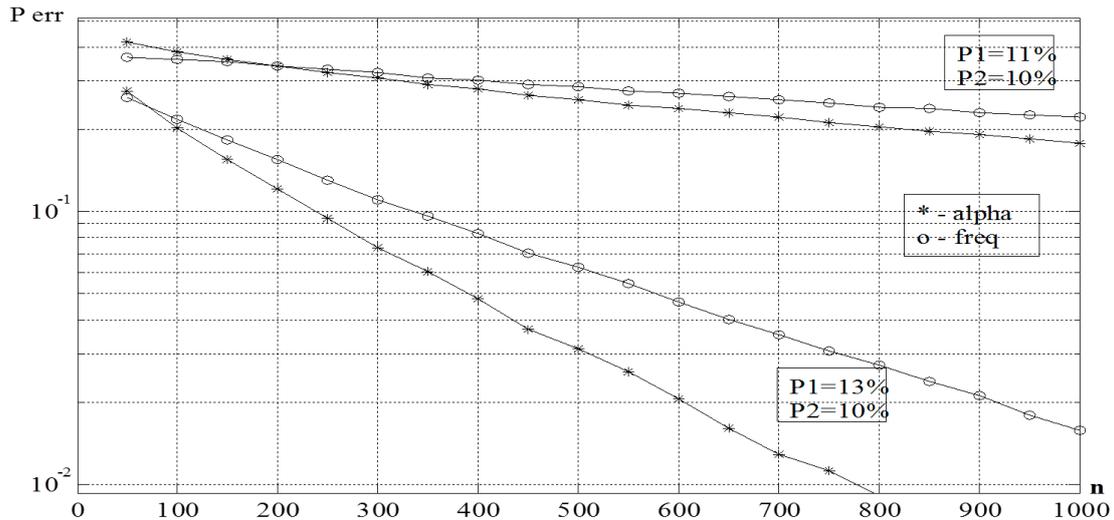


Рисунок 2. Зависимость вероятности ошибочного решения при сравнении САРГК от объема тестовой выборки звукозаписей

Для проверки предложенного метода проведен следующий эксперимент. Тестировались две системы, предназначенные для распознавания голосовых команд - названий цифр от 0 до 9. В первой системе использовалось семь гауссовых кривых для описания законов распределения в состояниях моделей скрытых марковских процессов, а во второй – 10, что обеспечивает более высокую точность акустических моделей звуков. Объем тестирующей выборки для каждой команды – 56.

Ошибки распознавания при тестировании отсутствовали, что не позволяет сравнить системы по относительной частоте ошибок распознавания. Использование же предложенного метода сравнения результатов распознавания показало, что, во-первых вторая система обеспечивает более высокую достоверность распознавания, а во-вторых, вероятность ошибки сравнения для некоторых команд достигала 49%, что свидетельствует о необходимости увеличить объем тестовой выборки.

В третьем разделе рассматриваются помехи, присутствующие на интервале длительности ключевого слова команды. Анализ влияния помехи и метода ее подавления целесообразно рассматривать как на уровне отдельных фрагментов сигнала, так и на уровне тестовых выборок голосовых команд, когда оцениваются усредненные характеристики изменений параметров РС, а также снижение достоверности распознавания.

Проанализирована возможность определения степени влияния помех на достоверность распознавания с помощью оценки $\hat{\alpha}$, рассмотренной в предыдущем разделе. Рассмотрим величину

$$\hat{b} = \frac{|\Phi(\hat{\alpha}_0) - \Phi(\hat{\alpha}_n)|}{0,5 - \Phi(\hat{\alpha}_0)}$$

Здесь $\hat{\alpha}_0$, $\hat{\alpha}_n$ - значения критерия $\hat{\alpha}$ для случаев отсутствия и наличия помехи.

Так как функция $\Phi(\alpha_n)$ Лапласа является монотонной, то разность, стоящая в числителе выражения, монотонно зависит от разности $\hat{\alpha}_0 - \hat{\alpha}_n$. Следовательно,

использование критерия $\hat{\alpha}$ при вычислении \hat{b} обеспечивает более точный учет влияния помех на достоверность распознавания по сравнению с использованием относительных частот.

В телефонной линии (ТЛ) можно выделить три основных типа помех: шум в полосе частот ТК; гармоническая помеха; импульсная помеха. Максимально допустимые искажения амплитудно-частотной характеристики (АЧХ) всего канала связи также нормируются.

Проведено экспериментальное исследование влияния указанных типов помех и искажений на достоверность распознавания двух команд: «ноль» и «один». В качестве системы параметров речевого сигнала использовались 12 кепстральных коэффициентов, логарифм энергии, а также первые две производные указанных параметров (всего 39 параметров). Размер окна анализа 25 мс, смещение кадров 10 мс, тип окна – окно Хэмминга. При использовании кепстральных коэффициентов влияние искажений АЧХ телефонного канала на параметры РС можно уменьшить, вычитая из вычисленных временных траекторий параметров их средние значения на интервале длительности ГК. Поэтому для повышения устойчивости системы распознавания к частотным искажениям использовались параметры с нулевым математическим ожиданием.

На вход САРГК подавались звукозаписи произнесений команд. К звукозаписям добавлялись имитации помех, характерных для телефонной линии. Были выделены основные виды помех в телефонном канале, а также их предельно допустимые величины согласно нормативным требованиям, предъявляемым к телефонному каналу:

- шум с равномерной спектральной плотностью в полосе 300-3400 Гц – минимальное отношение сигнал-шум 25 дБ;
- последовательность импульсов (помехи от источников питания), частота следования импульсов 100 Гц, минимальное отношение сигнал-шум 25 дБ;
- гармонические колебания (помехи от источников питания, вызывных сигналов). Частоты колебаний соответствовали границам частотного диапазона ТК и его середине (400, 1000, 2600 Гц). Минимальное отношение сигнал-шум для каждого колебания составило 32 дБ.

Анализировалось также влияние искажений амплитудно-частотной характеристики телефонного канала на достоверность распознавания. В Таблице 1 приведены результаты тестирования САРГК при наличии помех. В таблице представлены значения выборочного среднего $\hat{\mu}$, выборочного среднеквадратического отклонения $\hat{\sigma}$, а также величин: $\hat{\alpha}$, \hat{b} , $\hat{p}_{\text{ош}}$.

Название колонки «300-3400» соответствует звукозаписям команд без помех; «w_noise» - воздействие белого шума; «pulse» - воздействие импульсной помехи; «sin_left», «sin_middle» и «sin_right» - соответствуют воздействиям гармонических помех на частотах: 400, 1000 и 2600 Гц, соответственно; «fading» - соответствует частотным искажениям в телефонном канале.

Таблица 1. Значения критерия достоверности и относительной частоты ошибки распознавания при воздействии помех

Помеха	300-3400	w_noise	pulse	sin_left	sin_middle	sin_right	fading
«Ноль»							
$\hat{\mu}$	403,05	187,20	189,66	334,93	369,35	441,21	401,78
$\hat{\sigma}$	183,51	87,60	103,13	172,06	131,53	180,27	178,47
$\hat{\alpha}$	2,20	2,14	1,84	1,95	2,81	2,45	2,25
$\Phi(\hat{\alpha})$	0,4861	0,4838	0,4671	0,4744	0,4975	0,4931	0,4878
\hat{b}	0	0,1636	1,3652	0,8404	0,8218	0,4863	0,1208
$\hat{p}_{\text{ош}} (\%)$	0,00	0,91	1,09	0,91	0,00	0,00	0,00
«Один»							
$\hat{\mu}$	355,78	118,26	121,27	161,61	273,91	301,96	340,85
$\hat{\sigma}$	115,03	55,32	54,15	91,13	124,57	123,37	108,95
$\hat{\alpha}$	3,09	2,14	2,24	1,77	2,20	2,45	3,13
$\Phi(\hat{\alpha})$	0,499	0,4838	0,4875	0,4616	0,4861	0,4929	0,4991
\hat{b}	0	15,1647	11,5357	37,3336	12,8926	6,1372	0,1267
$\hat{p}_{\text{ош}} (\%)$	0,00	2,73	1,82	3,64	0,73	1,82	0,00

Из таблицы следует, что уменьшение числа ошибок распознавания сопровождается уменьшением отношения \hat{b} . Использование данного отношения становится особенно полезным, если число ошибок невелико или их вообще нет (при проведении тестирования на звуковой базе относительно небольшого объема), что не позволяет судить о значении вероятности ошибки с достаточной точностью.

Из таблицы также следует, что на результаты распознавания команды «Ноль» наибольшее влияние оказывает импульсная помеха. Следует отметить, что для данной команды воздействие среднечастотной гармонической помехи увеличивает достоверность распознавания относительно случая отсутствия помех. Использование относительной частоты $\hat{p}_{\text{ош}}$ ошибки распознавания в качестве оценки степени влияния помех на работу САРГК не позволяет обнаружить данный факт.

На результаты распознавания команды «Один» наибольшее влияние оказывает низкочастотная гармоническая помеха. В целом влияние помех на распознавание данной команды намного сильнее по сравнению с ГК «Ноль», что можно объяснить наличием маломощного звука «d» в голосовой команде «Один». Для обеих команд искажения АЧХ телефонной линии проявляют себя слабо, что обусловлено выбранной системой параметров речевого сигнала, которая подавляет эти искажения.

На промежуточных этапах проектирования САРГК целесообразно оценивать степень влияния помех на распознавание команд путем оценки изменений параметров огибающей спектра речевого сигнала (РС), которые являются исходными данными для алгоритма распознавания. Здесь целесообразно использовать синтезированные сегменты РС, что позволяет устранить влияние

вариаций формы сигнала возбуждения в электрической модели формирования РС на результаты изменения значений параметров РС под воздействием помех.

Проведен эксперимент по исследованию влияния белого шума и гармонических помех с частотами 400, 1000 и 2600 Гц на значения 10 коэффициентов частной корреляции (КЧК), соответствующих звуку «а0» (гласная «а» в ударном слоге) русской речи.

Определены зависимости средних значений КЧК от отношения сигнал-шум. Зависимости не монотонны, что затрудняет их аппроксимацию аналитическими выражениями. Воздействие гармонической помехи на КЧК снижается при совпадении ее частоты с частотой форманты.

С целью оценки эффективности метода подавления помех (на основе фильтра Винера), который используется в стандарте Европейского института стандартов в области телекоммуникаций ETSI ES 202 212 V1.1.2 (2007-01) для распределенных систем автоматического распознавания речи был проведен соответствующий эксперимент.

На речевые сигналы, которые соответствовали произнесениям ГК, накладывалась помеха, представляющая собой аддитивную смесь «розового» шума и периодической импульсной последовательности с частотой следования 390 Гц (запись с телефонной линии). В качестве голосовых команд использовались названия всех цифр, а также произнесения слов: «да», «нет» и «ошибка». Отношение сигнал-помеха при этом составило 25 дБ.

Для выборки голосовых команд без помехи, выборки с шумом и выборки, полученной на выходе фильтра, вычислялись кепстральные коэффициенты, а также их первые и вторые производные. Данные наборы параметров сравнивались между собой: вычислялось две дисперсии: разности параметров сигнала при отсутствии помехи и сигнала с помехой, а также разности параметров сигнала при отсутствии помехи и сигнала с отфильтрованной помехой. За счет фильтрации дисперсия разности снизилась на 39%.

Проведено контрольное распознавание голосовых команд с использованием фильтрации. В эксперименте использовалась тестовая выборка, состоящая из 3900 произнесений указанных выше команд, записанных шестью дикторами (1 женщина, 5 мужчин). Каждый диктор по 50 раз произнес каждую команду. Данная выборка была записана в лабораторных условиях. При этом отношение сигнал – помеха составило более 45 дБ.

При создании моделей ГК в качестве обучающей использовалась указанная выше тестовая выборка ГК. Для тестовой выборки без помехи частота ошибок распознавания составила 0,18%; для выборки с помехой - 39%; для выборки с помехой после фильтрации – 6,3%.

Недостатком рассмотренного выше метода подавления помехи является искажение спектра сигнала, обусловленное прохождением сигнала через фильтр.

В разделе рассмотрена также возможность подавления зашумленной периодической помехи с медленно меняющейся частотой методом компенсации. Предлагается следующий алгоритм ее подавления.

Детектор голосовой активности определяет интервал паузы, где нет сигнала. На интервале паузы определяется период T_p детерминированной составляющей $u(t)$. Затем из паузы формируется фрагмент, длительность которого равна целому

числу периодов детерминированной составляющей помехи. Данный фрагмент задерживается на целое число L периодов (сдвигается по оси времени), умножается на некоторый коэффициент W и вычитается из ГК, пораженной помехой – происходит взаимная компенсация помехи, присутствующей на интервале действия ГК, и фрагмента помехи в паузе. Затем фрагмент вновь сдвигается по оси времени и вновь вычитается из сигнала, пораженного помехой. Процессы сдвига и вычитания продолжаются до тех пор, пока сигнал на протяжении всей своей длительности не будет очищен от помехи.

Из теории адаптивной компенсации помех следует, что, если мгновенные мощности копии «паузной» помехи и помехи, присутствующей на интервале ГК, одинаковы, то для максимального подавления помехи (по критерию минимума среднего квадрата ошибки) необходимо «паузную» помеху, задержанную во времени, умножить на коэффициент

$$W_0 = \frac{\overline{\xi(t)\xi(t - LT_p)}}{\overline{\xi(t - LT_p)^2}} = \rho_\xi(LT_p).$$

Здесь $\xi(t), \xi(t - LT_p)$ - реализации помехи в моменты времени t и $(t - LT_p)$, где L – целое число, T_p – период детерминированной составляющей помехи, LT_p – величина задержки «паузной» помехи. Черта сверху означает процедуру усреднения по ансамблю реализаций. Так как $\overline{\xi^2(t - LT_p)} = \overline{\xi^2(t)}$, то $W_0 = \rho_\xi(LT_p)$, где $\rho_\xi(LT_p)$ - нормированная автокорреляционная функция $\rho_\xi(\tau)$ помехи при $\tau = LT_p$. Усредняя W_0 на периоде T_p , и считая, что шумовые процессы $n(t)$ и $n(t - LT_p)$ не коррелированы, получаем

$$W_{01} = \frac{P_u}{P_u + P_n} = \frac{q}{1 + q},$$

где $q = P_u/P_n$. Здесь $P_u = \langle u^2(t) \rangle$, P_n - средние мощности периодической составляющей и шума, соответственно. Средняя мощность помехи после компенсации

$$P_{\xi \text{ ост}} = (P_u + P_n) \left[1 - \left(\frac{P_u}{P_u + P_n} \right)^2 \right] = P_n \left(1 + \frac{q}{1 + q} \right).$$

Из полученного выражения следует, что при увеличении q от нуля до бесконечности остаточная мощность помехи меняется от P_n до $2P_n$. Мощность остаточной после компенсации помехи практически определяется лишь уровнем шума. Если периодическая составляющая помехи достаточно мощная, то величина W_{01} близка единице. Сравним мощности остаточной помехи при $W=W_{opt}$ и при $W=1$.

$$\lambda = \frac{P_{\xi \text{ ост } 1}}{P_{\xi \text{ ост}}} = 1 + \frac{1}{1 + 2q}.$$

При $q > 2$ величина $\lambda < 1,2$. То есть проигрыш в подавлении помехи при использовании $W=1$ вместо $W=W_{opt}$ очень мал и составляет менее 0,8дБ. Следовательно, целесообразно использовать $W=1$ и не определять значение $\rho_\xi(LT_p)$ автокорреляционной функции помехи.

Определим минимальное значение q , при котором компенсация обеспечивает достаточно глубокое подавление помех при $W=1$. Уровень подавления помехи

$$\gamma = \frac{P_{\xi}}{P_{\xi \text{ ост } 1}} = \frac{P_u + P_n}{2P_n} = 0,5(1 + q).$$

Если задаться величиной $\gamma > 2$, то величина q должна быть больше трех.

Рассмотрим теперь влияние погрешности в определении периода T_p на степень подавления помехи. Определим период как минимальное значение аргумента автокорреляционной функции помехи не равное нулю, при котором функция принимает максимальное значение. Имеют место два источника погрешностей при определении аргумента: дискретизация аргумента функции и влияние шума на результат определения максимума функции.

Шаг дискретизации равен периоду T_d дискретизации речевого сигнала. Максимальная погрешность определения аргумента равна половине периода дискретизации. Если помеха содержит гармоническое колебание с частотой f , то величина подавления помехи

$$\gamma = \frac{P_{\xi}}{P_{\xi \text{ ост } 1}} = \frac{P_u + P_n}{2P_n + 2P_u[1 - \cos(\pi f/f_d)]} = \frac{0,5(1 + q)}{1 + q[1 - \cos(\pi f/f_d)]}.$$

Если задаться величиной γ подавления помехи, то можно определить максимально допустимое отношение f/f_d .

Определим вероятность того, что за счет влияния шума оценка $\hat{\rho}_u(m\tau_0 + T_d)$ автокорреляционной функции станет больше оценки $\hat{\rho}_u(m\tau_0)$. В этом случае возникает погрешность определения периода равная периоду дискретизации. Если период T_p определяется по аргументам максимумов автокорреляционной функции помехи, то необходимо проанализировать значение случайной величины $\Delta\hat{R}(1) = \hat{R}(LB) - \hat{R}(LB + 1)$, где $\hat{R}(LB)$ и $\hat{R}(LB + 1)$ оценки автокорреляционной функции

$$\hat{R}(LB) = \frac{1}{N - LB} \sum_{i=LB}^{N-1} \xi(i) \xi(i - LB)$$

$$\hat{R}(LB + 1) = \frac{1}{N - LB - 1} \sum_{i=LB+1}^{N-1} \xi(i) \xi(i - LB - 1)$$

Здесь B – период детерминированной составляющей помехи, выраженный в виде числа периодов T_d дискретизации; L – порядковый номер рассматриваемого максимума автокорреляционной функции; N – длительность паузы, выраженная через число периодов T_d дискретизации. Величина $\xi(i)$ определяется суммой сигнала, помехи и шума.

Если случайная величина $\Delta\hat{R}(1) > 0$, то при значении дискретного времени равном LB оценка автокорреляционной функции периодической составляющей помехи принимает максимальное значение, и величина периода детерминированной составляющей помехи определяется без погрешности. Если же $\Delta\hat{R}(1) < 0$, то оценка автокорреляционной функции принимает максимальное значение при значении дискретного времени равном $LB+1$. В этом случае

величина периода детерминированной составляющей помехи определяется с погрешностью равной периоду дискретизации.

Если шум гауссов, то вероятность $p_{\text{ош}}$ того, что $\Delta\hat{R}(1) < 0$, определяется выражением

$$p_{\text{ош}} = 0,5 - \Phi(\alpha), \quad \alpha = \frac{\mu_{\Delta}}{\sqrt{D_{\Delta}}},$$

где $\Phi(\alpha)$ - функция Лапласа. Математическое ожидание μ_{Δ} случайной величины $\Delta\hat{R}(1)$ равно разности математических ожиданий случайных величин: $\hat{R}(LB), \hat{R}(LB + 1)$

$$\begin{aligned} \mu_{\Delta} &= \frac{1}{N-LB} \sum_{i=0}^{N-1-LB} u^2(i) - \frac{1}{N-LB-1} \sum_{i=0}^{N-2-LB} u(i)u(i+1) \approx \\ &\approx R_u(0) - R_u(1) \end{aligned}$$

Приближенное равенство тем точнее выполняется, чем в большей степени соблюдается условие $N-LB \gg LB$, и чем больше L . Величина дисперсии D_{Δ} случайной величины $\Delta\hat{R}(1)$ определяется выражением

$$\begin{aligned} D_{\Delta} &= \sigma_n^2 \sum_{i=0}^{N-LB-2} \left[\left(\frac{u(i+1)}{N-LB} - \frac{u(i)}{N-LB-1} \right)^2 + \left(\frac{u(i)}{N-LB} - \frac{u(i+1)}{N-LB-1} \right)^2 \right] + \\ &+ \sigma_n^2 \frac{u^2(N-1) + u^2(0)}{(N-LB)^2} + 2\sigma_n^4 \frac{(N-LB)^2 - (N-LB) + 1}{(N-LB)^2 * (N-LB-1)} \end{aligned}$$

Полученные выражения проверены с помощью моделирования.

Проанализированный алгоритм компенсации помех экспериментально проверен путем его использования для подавления импульсной помехи (сигнал «занято» в телефонной линии). Форма помехи: «пачки» длительностью 0,25с искаженных импульсов, следующих с частотой 0,3Гц. Частота следования импульсов внутри «пачки» - 500Гц. Кроме того, на речевой сигнал накладывались шум и слабая относительно импульсов гармоническая помеха с частотой 50Гц. Частота дискретизации составила 10кГц. Глубина квантования - 16 бит. Подавление помехи на интервале длительности первой «пачки» импульсов составило 27,3дБ.

В четвертом разделе рассмотрены возможности адаптивной компенсации широкополосных радиопомех, присутствующих на магистральном участке телефонного канала связи. В каналах с селективно-частотными замираниями применяют разнесенный прием, что позволяет осуществить адаптивную компенсацию помех. В общем случае, помехи поступают на приемные антенны совместно со своими отражениями от сооружений, расположенных вблизи станции связи. Поэтому компенсация помех осуществляется с помощью адаптивных трансверсальных фильтров (АТФ), которые устанавливаются в приемные тракты.

Сложность АТФ определяется числом отводов линии задержки (ЛЗ). Наименьшее их количество имеет место, когда отводы расположены неравномерно, и их расположение соответствует задержкам отражений помехи.

Важно определить зависимость степени компенсации от погрешности реализации требуемых задержек в АТФ. Проанализирована работа двухканального адаптивного компенсатора помех, в котором используется алгоритм компенсации Хоуэлса - Эпплбаума. Результатом работы алгоритма является вектор коэффициентов весового суммирования (КВС) напряжений с выводов линий задержек

$$\vec{W}_\infty = (E + \mu R_\Pi)^{-1} \vec{B},$$

где E - единичная матрица размерности $2M \times 2M$; $R_\Pi = \overline{\overline{X_\Pi^*(t) X_\Pi^T(t)}}$ - корреляционная матрица помех; $\overline{X_\Pi(t)}$; \vec{W} ; \vec{B} - векторы помеховых компонентов на выходах ЛЗ, коэффициентов весового суммирования и управления размерности $2M$ соответственно; μ - коэффициент усиления в петле обратной связи. M - число отводов линии задержки в АТФ.

Ситуация, наиболее трудная для подавления помех, характеризуется следующими положениями:

- помеха полностью перекрывает по спектру полезный сигнал;
- отражения помехи поступают на антенны с различных направлений от достаточно далеко разнесенных отражателей, что приводит к некоррелированности помеховых компонентов на соседних отводах ЛЗ;
- абсолютные величины погрешностей реализации относительных запаздываний в отводах ЛЗ максимальны, причем знак погрешности в трансверсальном фильтре ТФ1 противоположен знаку погрешности во втором фильтре ТФ2.

Считаем, что интервал корреляции широкополосного сигнала меньше расстояния между соседними отводами ЛЗ в составе КП (случай "больших задержек"). Поэтому сигнальные компоненты на выходах ЛЗ не коррелированы. Пространственное разнесение антенн обуславливает слабую корреляцию селективно-частотных замираний сигналов в каналах приема. При достаточно широкой полосе принимаемых сигналов и наличии селективно-частотных замираний корреляционную матрицу сигналов на входах компенсатора можно считать диагональной.

$$R_c = P_{co} \begin{bmatrix} E & 0 \\ 0 & E \end{bmatrix}$$

где P_{co} - мощность сигнала на входе ЛЗ; $E, 0$ - соответственно единичная и нулевая матрицы размерности $M \times M$. При указанных условиях корреляционная матрица помех

$$R_\Pi = P_{\Pi 0} \begin{bmatrix} \vec{F}_1^+ \vec{F}_1 E & \rho_\Delta \vec{F}_2 \vec{F}_1^+ \\ \rho_\Delta \vec{F}_1 \vec{F}_2^+ & \vec{F}_2^+ \vec{F}_2 E \end{bmatrix},$$

где $P_{\Pi 0} = \overline{U_m^2} / 2R_{\text{вх}}$ - мощность максимального по уровню отражения на входе ЛЗ; $R_{\text{вх}}$ - входное сопротивление ЛЗ; $U_m(t)$ - модуль комплексной огибающей максимального по уровню отражения (или прямого прохождения помехи) на входе ЛЗ; $\vec{F}_1 = \Phi_2 \vec{\theta}_1$; $\vec{F}_2 = \Phi_1 \vec{\theta}_2$; $\Phi_1 = \text{diag}[\exp(j\varphi_{ii}^{(1)})]$ - диагональная матрица

фазовых сдвигов на выводах линии задержки, обусловленных погрешностями реализации запаздываний отражений, взятых с противоположными знаками, для ТФ1; $\Phi_2 = \text{diag} [\exp(j\varphi_{ii}^{(2)})]$ – матрица фазовых сдвигов для ТФ2; $\vec{\theta}_1^T = [V_{11}, V_{12}, \dots, V_{1M}]$ – вектор коэффициентов ослаблений отражений для первого тракта распространения помехи; $\vec{\theta}_2^T = [V_{21}, V_{22}, \dots, V_{2M}]$ – вектор для второго тракта; $\rho_\Delta = |\rho(\Delta\tau)|$ – модуль нормированной корреляционной функции помехового излучения при временном сдвиге равном удвоенной абсолютной величине максимальной погрешности реализации относительных запаздываний.

Для максимального приближения векторов \vec{W}_1, \vec{W}_2 , которые относятся к адаптивным трансверсальным фильтрам АТФ, установленным в каналах приема, и которые составляют общий вектор \vec{W} , к оптимальным значениям, \vec{W}_{01} и \vec{W}_{02} целесообразно установить векторы управляющих коэффициентов в каналах приема $\vec{B}_1 = k_1 \vec{F}_2$; $\vec{B}_2 = k_2 \vec{F}_1$. Здесь k_1, k_2 – некоторые коэффициенты. Показано, что $k_2/k_1 \approx -1/\rho_\Delta$.

Проведенное моделирование адаптивного компенсатора помех подтверждает справедливость полученных выражений.

Использование адаптивных компенсаторов помех (АКП) наталкивается на опасность ошибочного подавления полезного сигнала за счет влияния его на работу цепей адаптации. Встает задача разработки алгоритмов адаптивной компенсации помех, обеспечивающих нечувствительность цепей адаптации к сигналу. Рассмотрена возможность уменьшения влияния фазоманипулированного сигнала, прошедшего канал с достаточно медленно меняющейся частотной характеристикой, на настройку АКП при подавлении помех с широкой полосой частот. Суть предлагаемого алгоритма анализируется на примере двухэлементного АКП, структура которого изображена на Рисунке 4.

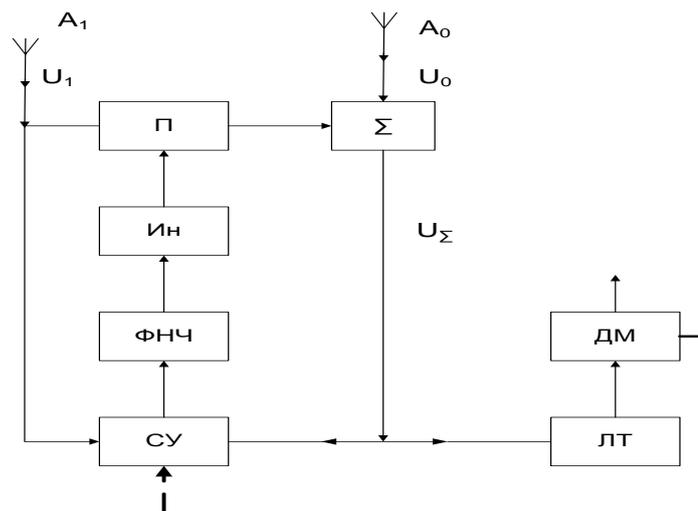


Рисунок 3. Структура двухвходового компенсатора помех с компенсацией влияния сигнала на цепи адаптации

Схема управления (СУ) определяет коэффициент W весового суммирования

сигнала с комплексной огибающей U_1 , поступающего от антенны A_1 , с сигналом комплексной амплитуды U_0 от антенны A_0 . Перемножитель (П) осуществляет операцию комплексного перемножения $U_1 * W$. Инвертор (Ин) и фильтр нижних частот (ФНЧ) служат для установки требуемого знака и реализации необходимого усреднения корреляционной обратной связи в АКП. К выходу сумматора (Σ) подключается линейный тракт (ЛТ) приемника с демодулятором (Дм) на выходе. Структура АКП и дальнейший ее анализ приведены для комплексных огибающих сигнала и помехи.

При реализации часто используемого алгоритма наименьших квадратов СУ содержит только перемножитель комплексных огибающих, выход которого подключен к указанному на схеме ФНЧ. Значение комплексного коэффициента W весового суммирования напряжений U_1 и U_0 определяется выражением:

$$W = -\mu \bar{y}, y = U_1^* U_\Sigma = U_1^* U_0 + W U_1^* U_1,$$

где: μ - коэффициент передачи петли корреляционной обратной связи, * - знак комплексного сопряжения, а черта сверху $\bar{}$ означает операцию усреднения во времени. Так как U_1 и U_Σ содержат сигнальный компонент, то величина W оказывается от него зависимой, что снижает степень подавления помехи.

Чтобы снизить эту зависимость в состав СУ дополнительно вводятся элемент задержки, перемножитель сигналов и алгебраический сумматор. Длительность задержки устанавливается кратной длительности информационной посылки и имеет величину большую интервала корреляции помехи. Через элемент задержки пропускается напряжение y , которое перемножается с напряжением U_1 . После усреднения в ФНЧ на выходе дополнительного перемножителя присутствует только постоянное напряжение, обусловленное присутствием сигнального компонента.

Напряжение меняет знак в зависимости от фазы передаваемых информационных посылок. Информация, получаемая с выхода демодулятора, позволяет корректировать знак напряжения так, чтобы оно имело полярность противоположную составляющей напряжения на выходе основного перемножителя, обусловленной влиянием сигнального компонента. При последующем суммировании выходных напряжений перемножителей происходит взаимная компенсация их составляющих, обусловленных наличием фазоманипулированного сигнала в составе напряжений y и U_1 .

Получены выражения, позволяющие определить зависимость выходного отношения сигнал-помеха на выходе компенсатора от аналогичного отношения на входе. Графики зависимостей для случая, когда внешняя помеха является белым шумом, представлены на Рисунке 4.

Графики построены для разных значений $q_{ш}$. Величина $q_{п}$ изменяется в пределах: 0...4. Здесь $q_{п} = \overline{|U_{п}|^2} / \overline{|U_{с}|^2}$; $q_{ш} = \overline{|U_{ш}|^2} / \overline{|U_{с}|^2}$. Соответствие графиков значениям $q_{ш}$ отражено Таблицей 2. Здесь в третьей и четвертых колонках указаны соответствующие значения сигнал-шум и вероятность ошибки демодулятора сигналов с относительной фазовой манипуляцией (ОФМ) при наличии только шума (внешняя помеха отсутствует).

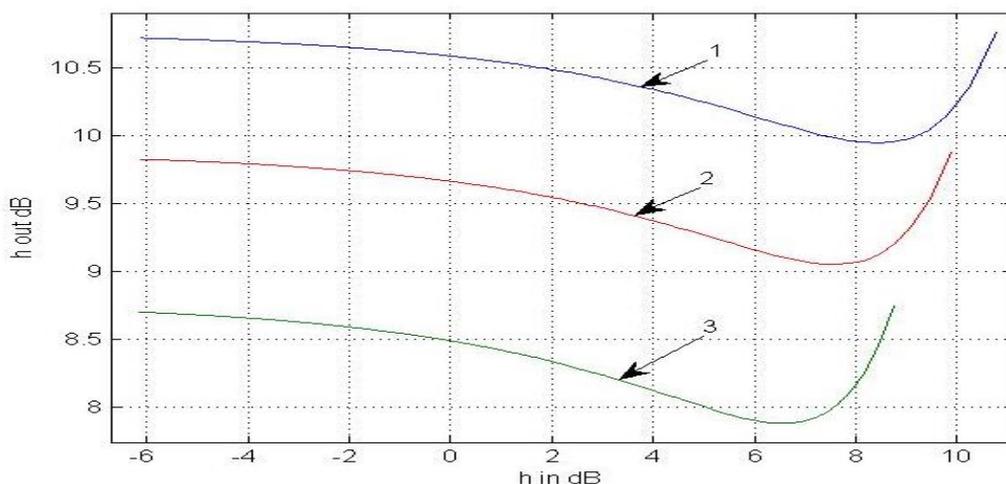


Рисунок 4. Графики зависимости отношения сигнал-шум на выходе компенсатора от отношения сигнал-шум на его входе

Таблица 2. Вероятность ошибки демодулятора ОФМ – сигнала в зависимости от отношения сигнал-шум

№ графика	$q_{ш}$	$h_{ш}$ дБ	$p_{офм}$
1	0,0839	10,76	10^{-6}
2	0,1028	9,88	10^{-5}
3	0,1327	8,77	10^{-4}

Из графиков следует, что при значительных изменениях величины $q_{ш}$ отношение h_{Σ} «сигнал/(помеха+шум)» меняется незначительно и имеет большое значение, которое обеспечивает малую вероятность ошибки демодулятора. С уменьшением уровня шума относительно сигнала возрастает отношение h_{Σ} , что обеспечивает меньшую вероятность ошибки. Минимальные значения h_{Σ} , наблюдаемые на графиках, соответствуют случаю, когда внешняя помеха становится соизмеримой с уровнем шума.

Использование алгоритма позволяет подавить помеху даже в случае, когда отношение сигнал-помеха на входе компенсатора больше единицы. В частности, при отношении сигнал-помеха на входе равном 4 дБ отношение сигнал-помеха на выходе равно 8,2 дБ.

В разделе описана также работа предложенного компенсатора импульсных помех.

В пятом разделе рассмотрен функциональный состав программных средств, используемых при проектировании САРГК. Рассмотрены разработанные программные средства, предназначенные для исследования помехоустойчивости САРГК. В частности, рассмотрены возможности исследования алгоритма адаптивной компенсации помех. Разработан программный комплекс с возможностью графического программирования различных видов исследования помехоустойчивости САРГК. Рассмотрена работа программы для оперативного редактирования сети акустических моделей.

Проанализирована работа программных средств, предназначенных для оценки достоверности распознавания команд. На Рисунке 5. представлен результат

работы программы - графики зависимостей значения предложенного критерия достоверности распознавания 10 команд (названий цифр) от отношения сигнал-шум.

Из графиков, в частности, следует, что в диапазоне (28...48 дБ) значений отношения сигнал-помеха наименьшая достоверность распознавания соответствует команде «три».

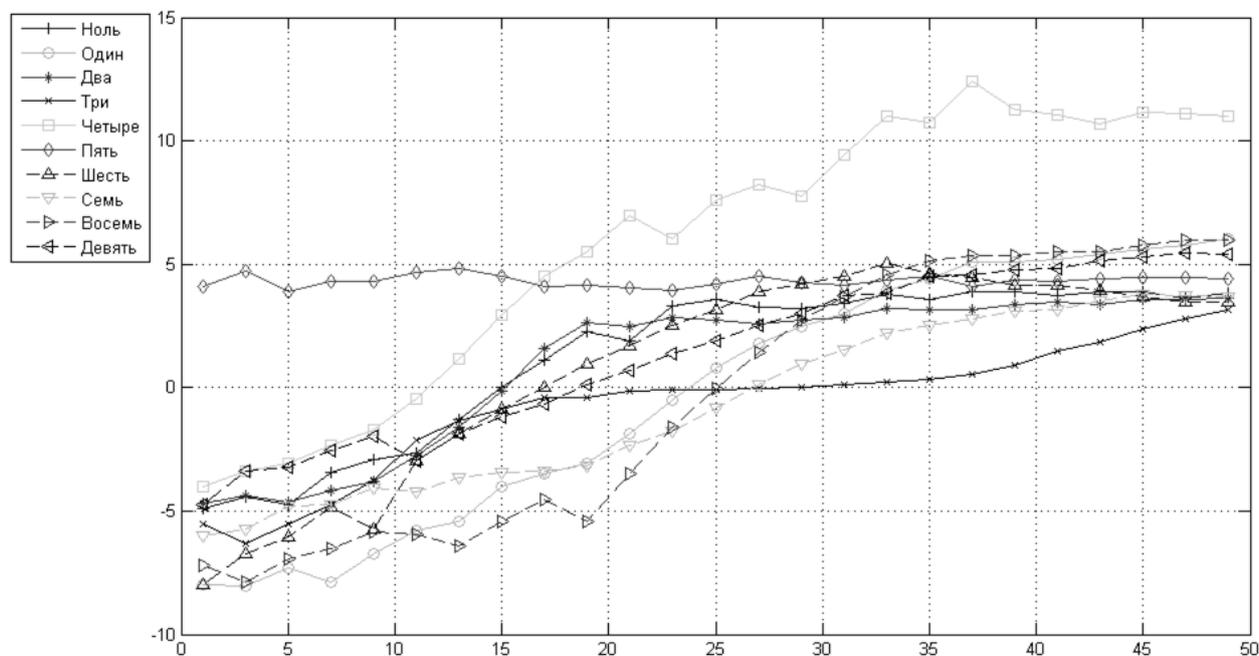


Рисунок 5. Графики зависимости значения критерия достоверности распознавания от отношения сигнал-шум

В шестом разделе рассмотрены аппаратные средства, необходимые для исследования помехоустойчивости САРГК. В частности, разработана система сбора информации об ошибках распознавания при организации прямого доступа к абоненту учрежденческой АТС. Структура системы представлена на Рисунке 6.

При поступлении вызова от абонента телефонной сети общего пользования (ТСОП) срабатывает детектор вызова, что обеспечивает срабатывание ключевой схемы (Ключ) – происходит занятие телефонной линии и начинается «диалог» системы с абонентом. Блок «Преобразование линии связи» преобразует двухпроводную линию связи в четырехпроводную, и появляется возможность использования подсистемы синтеза голосовых сообщений (Синтез голосовых сообщений), а также распознавания голосовых команд. Этот же блок обеспечивает гальваническую развязку телефонной линии и персонального компьютера, а также подавление местного эффекта. По результатам распознавания формируются сигналы управления блоками, указанными на схеме.

Если ошибки распознавания нет, то с помощью блока «Ключ» осуществляется набор номера вызываемого абонента УАТС. Далее САРГК в обслуживании телефонного вызова не используется. Если же произошла ошибка распознавания, то весь предыдущий «диалог» системы с абонентом записывается с указанием даты и времени события. Одновременно абоненту предлагается повторить команду. Если он отказывается от дальнейшего взаимодействия с

САРГК, то фиксируется либо повторная ошибка, либо факт превышения длительности паузы заданного порога. В этом случае САРГК отключается, и блок «Ключ» набирает номер телефонного аппарата (ТА) оператора, который обеспечивает обработку вызова.

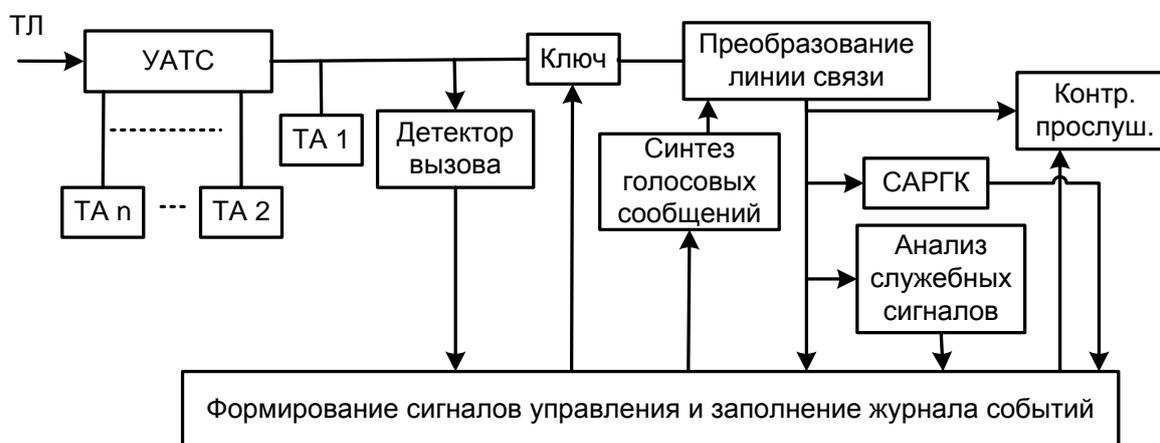


Рисунок 6. Схема организации прямого доступа с использованием САРГК

Блок «Анализ служебных сигналов» предназначен для обработки служебных сигналов типа «Отбой» в телефонной линии, когда, например, абонент ТСОП внезапно положит трубку. Блок «Контр. прослуш.» обеспечивает контрольное прослушивание «диалога» системы с абонентом для уточнения причины возникновения ошибки распознавания либо для ускорения процесса обслуживания телефонного вызова.

Соответствующее ПО устанавливается на компьютере работника учреждения. Телефонный аппарат работника обозначен на схеме как ТА 1. Указанный работник выполняет функции оператора системы. Если же работник отсутствует на рабочем месте, то функции оператора выполняет другой работник.

При разработке средств адаптивной компенсации широкополосных помех, рассмотренных в разделе 4, возникает задача определения степени подавления помех. Для решения данной задачи разработана схема двухканального устройства компенсационной обработки сигналов (УКОС).

В состав устройства входят управляемые линии задержки. Их количество определяется максимально возможным числом отражений помехи, на которое рассчитан компенсатор помех. На входы линий задержки поступают помехи. Выходы линий задержки подключены к входам многоканального адаптивного компенсатора помех, работающего по алгоритму Хоуэлса-Эплбаума.

Меняя величины задержек и вектор управляющих коэффициентов и контролируя уровень помехи на выходе компенсатора, можно оценить зависимость степени подавления погрешности реализации требуемых задержек и значений вектора управляющих коэффициентов.

Разработана схема испытаний компенсатора помех в составе станции связи, которая во время испытаний работает в режиме кольцевой проверки.

С целью сокращения затрат на аппаратуру, используемую при испытаниях, предложено автоматизированную перестройку частоты имитатора помех реализовать на основе цифровой системы частотной автоподстройки (ЦСЧАП). В состав системы входит электронно-счетный частотомер, выпускаемый промышленностью. Для повышения быстродействия ЦСЧАП и сокращения вычислительных затрат на ее реализацию предложены устройства, защищенные авторским свидетельством.

В Заключении приведены результаты решения поставленных в диссертации задач.

В Приложениях представлены результаты экспериментов по оценке вероятности ложного срабатывания, математические выкладки соответствующие анализу влияния шума на оценку периода помехи для осуществления ее подавления методом компенсации, акты внедрения. Кроме того, представлены листинги основных функций, разработанных в системе Matlab. Функции предназначены для моделирования адаптивного компенсатора помех, определения вероятности ошибки при сравнении результатов распознавания команд, сравнения результатов распознавания команд по помехоустойчивости.

ОСНОВНЫЕ РЕЗУЛЬТАТЫ И ВЫВОДЫ

1. Из-за высокой изменчивости произнесения голосовых команд проектирование САРГК требует проведения экспериментов – большого количества процедур тестирования системы объемными выборками звукозаписей для определения оптимальных параметров алгоритмов обработки речевых сигналов в условиях воздействия помех, а также для определения параметров достоверности распознавания команд.

2. Помехоустойчивость САРГК определяется степенью изменения параметров достоверности распознавания – степенью увеличения относительных частот ложного пропуска команды и ложного срабатывания системы под воздействием помех.

3. С расширением сферы использования САРГК увеличивается разнообразие помех и расширяется круг пользователей САРГК. Расширение круга пользователей требует повышения достоверности распознавания команд, что наряду с ростом разнообразия помех приводит к увеличению объема экспериментов при разработке САРГК.

4. Предложен метод сравнения результатов распознавания команды, полученных в нескольких процедурах тестирования, по достоверности распознавания. Метод позволяет сократить объем тестовой выборки по сравнению с известным методом. В частности, если вероятности ложного пропуска команды сравниваемых результатов равны 10% и 13%, то при вероятности ошибки сравнения 5% сокращение объема равно 31%.

5. Предложен метод оценки вероятности ложного срабатывания с учетом отличия произнесений посторонних слов от произнесения слова команды. Метод позволяет снизить объем тестирования системы при определении оценки для звукового материала, отличающегося от тестовой выборки звукозаписей. Объем

тестирования определяется лишь дополнительными звукозаписями, которые отсутствовали при тестировании на стадии проектирования САРГК.

6. На основе предложенного критерия достоверности распознавания разработана методика определения типа помехи, наиболее сильно влияющей на результаты распознавания.

7. Предложен алгоритм подавления зашумленной квазипериодической помехи. Получены выражения, позволяющие учесть влияние шума и выбора частоты дискретизации при определении периода помехи с целью последующего ее подавления методом неадаптивной компенсации.

8. Предложен алгоритм подавления влияния сигнала на работу цепей адаптации при компенсации широкополосных помех. Получены выражения, позволяющие определить отношение сигнал-помеха на выходе компенсатора от аналогичного отношения на его входе.

9. Разработаны программные и аппаратные средства для исследования помехоустойчивости САРГК.

ПУБЛИКАЦИИ, ОТРАЖАЮЩИЕ ОСНОВНЫЕ РЕЗУЛЬТАТЫ ДИССЕРТАЦИИ

Статьи в рецензируемых журналах, рекомендованных ВАК для публикации результатов диссертаций

1. Левин, Е.К. Адаптивная компенсация широкополосных радиопомех / Е.К. Левин // Проектирование и технология электронных средств. – 2013. - №1. - С. 2-6.
2. Левин, Е.К. Оценка вероятности ложного срабатывания системы автоматического распознавания голосовых команд / Е.К. Левин, П.С. Рагузин, М.Ю. Татарникова // Научно-технические ведомости СПбГПУ. Информатика, телекоммуникации, управление – 2011. - № 5 – С. 7 – 13.
3. Левин, Е.К. Компенсация помех при автоматическом распознавании голосовых команд в телефонии / Е.К. Левин // Проектирование и технология электронных средств. – 2011. – №3. – С.45-49.
4. Левин, Е.К. Оценка помехоустойчивости автоматического распознавания голосовых команд в телефонии / Е.К. Левин, О.Р. Никитин // Радиотехнические и телекоммуникационные системы – 2011. - №4 – С. 75-80.
5. Левин, Е.К. Уменьшение объема звуковой базы, используемой при тестировании систем автоматического распознавания голосовых команд / Е.К. Левин // Проектирование и технология электронных средств. – 2010. - №2. - С. 42-44.
6. Левин, Е.К. Оценка вероятности ложного срабатывания системы автоматического распознавания голосовых команд / Е.К. Левин, П.С. Рагузин // Проектирование и технология электронных средств.- 2009.- №4. - С.57-61.

7. Левин, Е.К. Экспериментальные исследования системы распознавания голосовых команд / Е.К. Левин, К.Е. Левин, О.Р. Никитин // Проектирование и технология электронных средств. – 2005. – № 3. – С.38-44.
8. Левин, Е.К. Использование программного моделирующего комплекса для исследования точности оценки огибающей спектра речевого сигнала / Е.К. Левин, К.Е. Левин, О.Р. Никитин // Проектирование и технология электронных средств. – 2004. – №1. – С.49-53.
9. Левин, Е.К. Моделирование процесса создания эталонов голосовых команд для систем автоматического распознавания речи / Е.К. Левин, К. Е. Левин // Цифровая обработка сигналов. – 2002. - №2. - С. 21-22.
10. Левин, Е.К. Использование голосового управления в телефонной справочной системе медицинского учреждения / Е.К. Левин, О.Р. Никитин // Биомедицинская радиоэлектроника. – 2000. – №6. – С.22-25.
11. Левин, Е.К. Адаптивная компенсация помех с использованием обратной связи по решению /Е.К. Левин // Электросвязь, 1998. - № 3. - С. 32 - 34.
12. Левин, Е.К. Метод подавления совокупности помех от источника излучения и пассивных переизлучателей / Е.К. Левин, А.С. Немировский // Электросвязь, 1987. - № 9. - С. 48 - 51.
13. Коньков, В.А. Цифровой измеритель параметров амплитудно-частотных характеристик четырехполюсников / В.А. Коньков, А.Б. Казаринов, Е.К. Левин, А.Д. Поздняков // Приборы и техника эксперимента. – 1976. - № 2. – С.253.

Авторские свидетельства на изобретения

1. Устройство приема широкополосных сигналов с двукратным разнесением : а.с. 1286079 СССР : МКИЗ Н 04 В 7/04 /Е.К. Левин, И.М. Покровская, П.А. Полушин, А.Г. Самойлов (СССР). - №3880798/24-09; заявл. 05.04.85; опубл. 15.02.87 Бюл. № 3. – 14 с. : ил.
2. Устройство приема сигналов с двукратным разнесением : а.с. 1092741 СССР : МКИЗ Н 04 В 7/02; Н 04 В 7/08 /Е.К. Левин, А.А. Покровский, П.А. Полушин, А.Г. Самойлов (СССР). - №3534593/18-09; заявл. 06.01.83 ; опубл. 15.05.84 Бюл. № 18. – 8 с. : ил.
3. Цифровое устройство частотной автоподстройки : а.с. № 1104659 СССР : МКИЗ Н 03 L 7/00 / Е.К. Левин (СССР) №3539904/18-09; заявл. 11.01.83, опубл. 23.07.84 Бюл. № 27. – 8 с. : ил.
4. Устройство для автоматической подстройки частоты : а.с. № 511669: МКИ2 Н 03 В 3/04 / Е.К. Левин (СССР) №2072577/09; заявл. 04.11.74, опубл. 25.04.76 Бюл. № 15. – 2 с. : ил.

Монография, научные статьи и материалы конференций

1. Левин, Е.К. Средства исследования помехоустойчивости систем распознавания голосовых команд в телефонии. Монография. / Е.К. Левин – Владимир: Рост, 2014. – 234с.
2. Левин, Е.К. Зависимость ложного срабатывания системы распознавания голосовых команд от параметров произнесений посторонних слов. / Е.К. Левин, П.С. Рагузин // Перспективные технологии в средствах передачи информации. Материалы докладов девятой международной научно-технической конференции. – / Владим. гос. университет – Владимир: ВлГУ. т.1- 2011. – С.185 - 188.
3. Левин, Е.К. Экспериментальное исследование системы автоматического распознавания голосовых команд при наличии произнесений посторонних слов / Е.К. Левин // Перспективные технологии в средствах передачи информации. Материалы 8-ой международной НТК. Т.2. / Владим. гос. ун-т – Владимир, 2009.- С.156 – 158.
4. Левин, Е.К. О тестировании систем автоматического распознавания голосовых команд / Е.К. Левин // 8-я Международная НТК «Физика и радиоэлектроника в медицине и экологии», 2008г. - Владимир. - Книга 2 – С.140-145.
5. Левин, Е.К. Сравнение телефонных систем автоматического распознавания голосовых команд по достоверности распознавания / Е.К. Левин, К. Е. Левин // Перспективные технологии в средствах передачи информации. Материалы 7-ой международной научно-технической конференции. – / Владим. гос. университет – Владимир: РОСТ, 2007. – С.70 - 73.
6. Левин, К.Е. Система прямого доступа к абонентам АТС учреждения, управляемая голосом // Е.К. Левин, К.Е. Левин, О.Р.Никитин // Методы и устройства передачи и обработки информации. Межвуз. Сб. научн. тр. - Вып.7/ СПб.: Гидрометеиздат, 2006. - С.223-229.
7. Levin, E. K. The Noise Suppression for Automatic Speech Commands Recognition [Text] / E. K Levin, K.E. Levin // SPECOM'2005. – University of Patras, Patras, Greece, - 2005. - P. 507-509.
8. Левин, Е.К. Система для исследования алгоритма распознавания речевых сигналов / Е.К. Левин, К.Е. Левин // Биомедицинские технологии и радиоэлектроника. – 2004. – №3. – С.28-32.
9. Левин, Е.К. Исследование влияния аддитивного шума на точность определения параметров речевых сигналов при их распознавании / Е.К. Левин, К.Е. Левин // Физика и радиоэлектроника в медицине и экологии Материалы докладов шестой международной научно-технической конференции. В 2 т. – Владимир, 2004. – Т. 1. С. 148-150.
10. Левин, Е.К. Влияние помех в телефонном канале связи на точность автоматического распознавания голосовых команд / Е.К. Левин, К.Е. Левин // Обработка сигналов в системах телефонной связи и вещания. Материалы тринадцатой межрегиональной конференции. – М.: МНТОРЭС им.А.С. Попова, 2004. – С.90 - 92.

11. Levin, E. K. Using a simulation program complex for researches of speech recognition devices [Text] / E. K Levin, K.E. Levin // SPECOM'2003. – Moscow State Linguistic University, Moscow, Russia, - 2003. – P.216-219.
12. Левин, Е.К. Моделирование устройств обработки речевых сигналов в телефонии / Е.К. Левин // Перспективные технологии в средствах передачи информации. Материалы пятой международной научно-технической конференции. – / Владим. гос. университет – Владимир: ВлГУ. - 2003. – С.102 - 103.
13. Левин, Е.К. Моделирование системы автоматического распознавания голосовых команд / Е.К. Левин, К.Е. Левин // Обработка сигналов в системах телефонной связи и вещания. Материалы докладов одиннадцатой межрегиональной конференции. – М., МТУСИ 2002. – С.63-65.
14. Левин, Е.К. Исследование точности спектральной оценки речевого сигнала для автоответчика с голосовым управлением в медицинском учреждении / Е.К. Левин, К.Е. Левин // Физика и радиоэлектроника в медицине и экологии. Материалы докладов пятой международной научно-технической конференции. – Владимир, 2002. – С.146-147.
15. Левин, Е.К. Исследование процесса предварительной обработки речевых сигналов в системах автоматического распознавания голосовых команд / Е.К. Левин, К.Е. Левин // Перспективные технологии в средствах передачи информации. Материалы докладов четвертой международной научно-технической конференции. – Владимир, 2001. – С.129-130.
16. Левин, Е.К. Система распознавания голосовых команд / Е.К. Левин // Обработка сигналов в системах телефонной связи и вещания. Десятая межрегиональная конференция. Тезисы докладов. – М.: МНТОРЭС им.А.С. Попова, 2000. – С.121 - 125.
17. Левин, Е.К. Система распознавания голосовых команд / Е.К. Левин // Обработка сигналов в системах телефонной связи и вещания. Десятая межрегиональная конференция. Тезисы докладов. – М.: МНТОРЭС им.А.С. Попова, 2000. – С.121 - 125.
18. Левин, Е.К. Разработка телефонной справочной системы с голосовым управлением для медицинского учреждения / Е.К. Левин, О.Р. Никитин // Физика и радиоэлектроника в медицине и экологии – Материалы докладов четвертой международной научно-технической конференции. В 2 т. Т. 2. – Владимир, 2000. – С. 118-120.
19. Левин, Е.К. Исследование изменчивости параметров речевых команд / Е.К. Левин // Обработка сигналов в системах двусторонней телефонной связи. Восьмая межрегиональная конференция. Тезисы докладов. – М.: МНТОРЭС им.А.С. Попова, 1998. – С.83 - 84.
20. Левин, Е.К. Анализ измерителя основного тона речевого кодека на основе линейного предсказания / Е.К. Левин // Перспективные технологии в средствах передачи информации. Материалы второй международной научно-технической конференции. – / Владим. гос. университет – Владимир: ВлГУ. т.1- 1997. – С.225 - 227.

21. Левин, Е.К. Адаптивная компенсация помех в системах связи с сигналами, подверженными селективно-частотным замираниям / Е.К. Левин. // Методы и устройства первичной обработки сигналов в радиотехнических системах: Межвуз. сб. научн. трудов. г. Горький, Горьк. политехн. ин-т., 1988. - С.76-81.
22. Левин, Е.К. Флуктуации коэффициентов весового суммирования в устройствах адаптивной компенсации помех / Е.К.Левин.- М.,1987. - 12с. Деп. в ЦНТИ "Информсвязь" 20.08.87, №1152-св.87.
23. Левин, Е.К.. Анализ быстродействующей цифровой системы частотной автоподстройки / Е.К. Левин. - М., 1978. – 10 с. - Деп. в ЦНТИ Информсвязь 23.04.1978, №9.

Подписано в печать 17.07.14.

Формат 60×84/16. Усл. печ. л. 2,09. Тираж 100 экз.

Заказ

Издательство

Владимирского государственного университета
имени Александра Григорьевича и Николая Григорьевича Столетовых.
600000, Владимир, ул. Горького, 87.