

На правах рукописи



Лукьянов Дмитрий Игоревич

**Алгоритмы обработки речевых сигналов и классификации
психоэмоционального состояния человека на выходе каналов
передачи речевой информации при действии мешающих
факторов**

Специальности:

05.12.04 – «Радиотехника, в том числе системы и устройства телевидения» и

05.11.17 – «Приборы, системы и изделия медицинского назначения»

АВТОРЕФЕРАТ

диссертации на соискание ученой степени

кандидата технических наук

Рязань 2015

Работа выполнена в ФГБОУ ВПО
«Рязанский государственный радиотехнический университет»

- Научный руководитель - Кириллов Сергей Николаевич
заслуженный работник Высшей школы РФ,
доктор технических наук, профессор,
заведующий кафедрой Радиоправления и связи
ФГБОУ ВПО «Рязанский государственный
радиотехнический университет»
- Официальные оппоненты - Харламов Александр Александрович,
доктор технических наук, Федеральное
государственное бюджетное учреждение науки
Институт высшей нервной деятельности и
нейрофизиологии РАН, старший научный сотрудник
- Ромашкин Юрий Николаевич,
кандидат технических наук, Федеральное
государственное казенное учреждение «Войсковая
часть 35533, ведущий научный сотрудник
- Ведущая организация - ГКОУ ВПО «Академия Федеральной службы
охраны Российской Федерации», г.Орел

Защита состоится «25» июня 2015 г. в 11 часов на заседании
диссертационного совета Д 212.211.04 в ФГБОУ ВПО «Рязанский
государственный радиотехнический университет» по адресу: 390005, г. Рязань,
ул. Гагарина, д. 59/1.

С диссертацией можно ознакомиться в библиотеке ФГБОУ ВПО
«Рязанский государственный радиотехнический университет».

Автореферат разослан « » апреля 2015 г.

Ученый секретарь
диссертационного совета



Овечкин Г.В.

ОБЩАЯ ХАРАКТЕРИСТИКА ДИССЕРТАЦИОННОЙ РАБОТЫ

Актуальность темы. Разработка алгоритмов классификации психоэмоционального состояния (ПЭС) человека является одной из наиболее сложных задач в области обработки речевой информации.

Алгоритмы классификации ПЭС целесообразно применять в устройствах постобработки РС, а также системах передачи и отображения медико-биологической информации. К данному классу устройств можно отнести средства интерактивного взаимодействия человека и электронно-вычислительных машин (ЭВМ), криминалистической фоноскопии, медицинской диагностики, в информационно управляющих комплексах и др. Кроме того, большой интерес к системам удаленной оценки ПЭС человека может проявлять психосоматическое направление в медицине, позволяющее посредством анализа эмоций повысить точность диагностирования таких заболеваний, как бронхиальная астма, синдром раздраженной толстой кишки, эссенциальная артериальная гипертензия, головная боль, напряжения, головокружения и прочее.

В настоящее время существует несколько способов выявления ПЭС человека: посредством электрического сопротивления кожи и сердечно-сосудистой активности человека; с использованием мимики лица и движений тела; посредством обработки речевого сигнала (РС).

Анализ ПЭС человека посредством электрического сопротивления кожи и сердечно-сосудистой активности позволяет с достаточной точностью определять состояние испытуемого, в том числе осуществить оценку достоверности слов человека. К существенному недостатку этого метода можно отнести необходимость использования контактных датчиков, прикрепляемых к телу человека, что не всегда является возможным.

В последние годы использование мимики лица и движений тела для оценки ПЭС человека получило широкое распространение, что привело к серьезному развитию классификационных моделей. В то же время, данный метод также имеет ряд ограничений, а именно: необходимость анфасного расположения лица испытуемого, высокое качество видеоизображения, широкая полоса частот и высокая скорость передачи информации при осуществлении удаленной оценки ПЭС человека.

В случае обработки РС анализ ПЭС человека позволяет избежать недостатков, присущих двум другим методам, так как для функционирования этого метода необходим только микрофон.

В ряде основных проблем, существующих при разработке алгоритмов классификации ПЭС человека на основе РС, наиболее актуальными являются:

- 1 Создание объективной базы РС, имеющих эмоциональный окрас. Сложность решения этой проблемы обуславливается большим разнообразием разговорных языков, каждый из которых характеризуется своей эмоциональной окраской, а также спонтанностью проявления реальных ПЭС человека. Кроме того, существуют морально-этические аспекты создания объективной речевой базы ПЭС человека.

2 Разработка единой модели классификации ПЭС человека. Эта проблема является актуальной для всех известных методов оценки ПЭС человека, так как в настоящее время отсутствует точная терминология и описание эмоциональной окраски речи человека.

3 Определение значимых параметров РС, которые могут однозначно характеризовать ту или иную эмоцию, и при этом были бы нечувствительны к другим составляющим речи: смысловому содержанию фразы, индивидуальным особенностям голоса, биофизическим характеристикам диктора.

В настоящее время возрос интерес к определению ПЭС человека при анализе РС на выходе каналов передачи информации (КПИ) По этой причине требуется проведение отдельных исследований оценки влияния КПИ на эмоциональную окраску РС с целью определения вероятности правильной классификации ПЭС человека. Необходимо отметить, что на РС в КПИ действуют различные мешающие факторы в виде акустических помех, искажений, а также шумов аппаратуры.

Таким образом, актуальны задачи обоснования параметров классификации ПЭС человека на основе РС, исследования влияния КПИ на эмоциональную окраску РС, а также разработка алгоритма классификации ПЭС человека на основе анализа РС на выходе КПИ при действии мешающих факторов.

Степень разработанности темы. Впервые вопрос об оценке ПЭС человека возник в 1872 году в результате исследований Ч. Дарвина. Впоследствии интерес к изучению этого вопроса усиливался из года в год. Дальнейшее его развитие позволило получить первые модели, позволяющие классифицировать ПЭС человека посредством анализа РС. Предложенные модели характеризовались низкой вероятностью правильной классификации ПЭС, а также малым числом анализируемых эмоций, что обуславливалось низкими характеристиками доступных на тот момент средств обработки РС.

В настоящее время, проведено большое количество исследований в данной области. Наиболее существенный вклад в развитие методов оценки ПЭС человека посредством анализа РС внесли работы отечественных ученых В.П. Морозова, Е.П. Ильина, Э.А. Нушикян, Т.Н. Ушаковой, Н.И. Жинкина, Ю.Н. Ромашкина, а также работы зарубежных исследователей В. Вундта, К. Шерера, О. Фрая, К. Изарда, Ф. Счмийдера.

Цель и задачи работы. Основной целью диссертационной работы является разработка алгоритма классификации ПЭС человека на основе анализа РС на выходе КПИ при действии мешающих факторов.

Поставленная цель работы требует решения следующих задач:

1 Обоснования модели классификации ПЭС человека на основе обобщенной эмоциональной плоскости.

2 Создания базы РС, имеющих эмоциональный окрас, для обеспечения устойчивости алгоритмов определения значимых для оценки ПЭС человека параметров.

3 Определения требуемой полосы частот КПИ, содержащей эмоциональную составляющую РС.

4 Исследования влияния характеристик современных КПИ и мешающих

факторов на эмоциональную составляющую РС в интересах определения возможности оценки ПЭС человека на выходе КПИ.

5 Исследования влияния параметров алгоритмов асинхронного маскирования речи (АМР) на эмоциональную составляющую РС.

6 Обоснования структурной схемы алгоритма классификации ПЭС человека на основе анализа РС на выходе КПИ.

7 Определения требуемых вычислительных мощностей для реализации алгоритма классификации ПЭС человека на выходе КПИ.

Научная новизна. В рамках диссертационной работы были получены следующие новые научные результаты:

1 Определены новые значимые параметры РС, позволяющие классифицировать ПЭС человека, которыми являются коэффициенты линейного предсказания (ЛП), полученные на основе автокорреляционного метода, коэффициенты импульсной характеристики прямого фильтра и параметры фазовых портретов (ФП) РС.

2 Предложено использование метода группового учета аргументов (МГУА) для оценки значимости параметров РС в случае классификации ПЭС человека.

3 Предложено использование модифицированной фонетической функции речи А.А. Пирогова для оценки значимости статических и динамических параметров РС с точки зрения классификации ПЭС человека.

4 Проведено исследование влияния акустических помех на эмоциональную составляющую РС.

5 Проведены исследования представления Хургина-Яковлева при использовании производных первого – третьего порядка и показана эффективность представления при передаче эмоциональной составляющей РС.

6 Проведены исследования влияния характеристик КПИ на эмоциональную составляющую РС, показывающие необходимость их учета при классификации ПЭС человека.

7 Проведены исследования влияния алгоритмов АМР на эмоциональную составляющую РС.

8 Разработан алгоритм классификации ПЭС человека на основе анализа РС на выходе КПИ, учитывающий особенности передающего тракта и других мешающих факторов.

Теоретическая и практическая значимость работы. Представленные в работе алгоритмы классификации ПЭС человека на основе анализа РС могут быть использованы в таких радиотехнических устройствах, как системы передачи и хранения речевой информации, медицинские системы диагностики эмоционального состояния человека, системы голосового управления, информационно-справочные системы, системы автоматической идентификации и верификации информанта по голосу, системы криминалистической фоноскопии, системы конфиденциального доступа и закрытия речевой информации, а также в медицинских системах диагностики эмоционального состояния человека, системах открытого образования для лиц с ограниченными возможностями здоровья и т.д. Результаты диссертационной работы нашли

применение в разработках открытого акционерного общества «Российские космические системы» (ОАО «РКС») и учебном процессе ФГБОУ ВПО «РГРТУ», что подтверждено соответствующими актами.

Методология и методы исследования. В работе использовались методы статистической радиотехники и математической статистики, распознавания образов, динамического программирования, вычислительной математики, а также новые достижения в области цифровой обработки речевой информации. Данные теоретические методы сочетались с экспериментальными исследованиями на основе имитационного моделирования разработанных алгоритмов.

Положения, выносимые на защиту.

1 Применение коэффициентов ЛП, полученных на основе автокорреляционного метода; коэффициентов импульсной характеристики прямого фильтра и параметров ФП РС позволяет повысить вероятность правильной классификации ПЭС человека при том же числе параметров в среднем на 0,08 или снизить необходимое число параметров в 1,7 раза по сравнению с известными алгоритмами.

2 Алгоритм классификации ПЭС человека на основе анализа РС на выходе КПИ, обеспечивающий вероятность правильной классификации ПЭС человека по шкале «удовольствие-неудовольствие» до 0,93 и по шкале «расслабление-напряжение» до 0,83.

3 Алгоритм снижения влияния акустических помех и защиты РС на основе представления Хургина-Яковлева, обеспечивающий выигрыш в качестве восстановленного сигнала в среднем на 0,5 балла согласно ГОСТ Р 50840-95 и увеличивающий вероятность правильной классификации ПЭС в среднем на 0,04 при действии акустических помех.

Достоверность. Достоверность результатов и выводов, полученных в диссертационной работе, обосновывается корректностью численных экспериментов, качественным и количественным сопоставлением с известными положениями теории обработки и восприятия РС.

Апробация работы. Результаты работы докладывались на следующих конференциях:

1 Семнадцатая международная научно-техническая конференция «Проблемы передачи и обработки информации в сетях и системах телекоммуникаций». Рязань, 2012.

2 Двадцать четвертая всероссийская научно-техническая конференция студентов, молодых ученых и специалистов «Биотехнические, медицинские и экологические системы и комплексы». Рязань, 2012.

3 Пятьдесят девятая студенческая научно-техническая конференция. Рязань, 2012.

4 Семнадцатая всероссийская научно-техническая конференция студентов, молодых ученых и специалистов. Рязань, 2012.

5 Вторая итоговая конференция молодых инноваторов по программе «Умник» в Рязанской области. Рязань, 2012.

6 Двадцать пятая всероссийская научно-техническая конференция

студентов, молодых ученых и специалистов «Биотехнические, медицинские и экологические системы и комплексы». Рязань, 2012.

7 Десятая международная научная конференция «Перспективные технологии в средствах передачи информации». Владимир, 2013.

8 Шестая международная научно-техническая конференция «Космонавтика. Радиоэлектроника. Геоинформатика». Рязань, 2013.

9 Восемнадцатая всероссийская научно-техническая конференция студентов, молодых ученых и специалистов «Новые информационные технологии в научных исследованиях». Рязань, 2013.

10 Шестнадцатая международная конференция «Цифровая обработка сигналов и ее применение – DSPA-2014». Москва, 2014.

11 Семнадцатая международная телекоммуникационная конференция молодых ученых и студентов «Молодежь и наука». Москва, 2014.

12 Международная конференция «Радиоэлектронные устройства и системы для инфокоммуникационных технологий – РЭУС-2014». Москва, 2014.

13 Четвертая всероссийская научно-техническая конференция «Информационно-измерительные и управляющие системы военной техники». Владимир, 2014.

14 Девятнадцатая всероссийская научно-техническая конференция студентов, молодых ученых и специалистов «Новые информационные технологии в научных исследованиях». Рязань, 2014.

15 Двадцать седьмая всероссийская научно-техническая конференция студентов, молодых ученых и специалистов «Биотехнические, медицинские и экологические системы и комплексы». Рязань, 2014.

16 Восемнадцатая международная телекоммуникационная конференция молодых ученых и студентов «Молодежь и наука». Москва, 2015.

Публикации. По теме диссертации опубликовано 20 работ: 3 статьи в научно-технических журналах, рекомендованных ВАК, 1 статья в межвузовском сборнике трудов и 16 тезисов докладов на конференциях.

Структура и объем работы. Диссертационная работа состоит из введения, трех глав, заключения, списка литературы из 179 наименований и одного приложения. Диссертация изложена на 180 страницах, из которых 133 страницы основного текста, 26 таблиц и 70 рисунков.

ОСНОВНОЕ СОДЕРЖАНИЕ ДИССЕРТАЦИОННОЙ РАБОТЫ

Во введении обоснована актуальность выбранной темы, определены цель и решаемые в работе задачи. Изложены новые научные результаты, полученные в работе, показаны ее практическая ценность и апробация. Сформулированы основные положения, выносимые на защиту.

В первой главе произведен анализ влияния психоэмоционального состояния человека на параметры РС.

Обоснована модель классификации ПЭС человека. Предложена модель эмоциональной плоскости (рисунок 1), являющаяся модификацией модели В. Вундта. Преимущество предложенной модели заключается в меньшем

количестве подплоскостей при неограниченном числе охватываемых эмоциональных состояний. Исходя из широкого распространения таких терминов эмоций, как «радость», «ликование», «страх», «гнев» и др., целесообразно разместить их на плоскости, что позволяет существенно облегчить субъективное восприятие ПЭС аудиторами, при этом не исключая возможность автоматической классификации ПЭС.

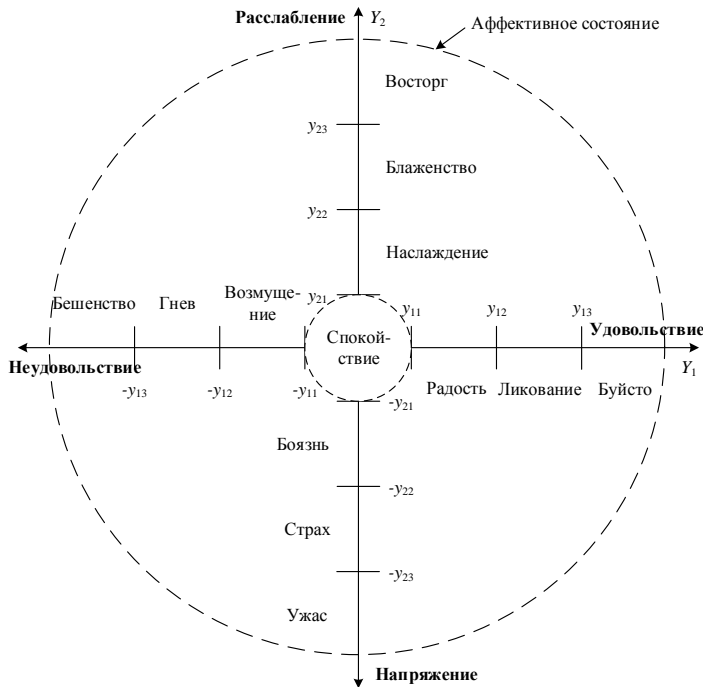


Рисунок 1

Собрана и обоснована база речевых материалов, записанная на русском языке, которая включала в себя два набора эмоций: спровоцированных и естественных. При этом в базе отсутствовали актерские эмоции, что объясняется не полной корреляцией результатов анализа этого типа эмоций с реальными.

Первый набор включал в себя фразы, отражающие спровоцированные эмоции, с участием 12 дикторов (6 мужчин и 6 женщин). Дикторы, участвовавшие в подготовке этого набора, не знали целей

исследования до момента окончания сбора записей, что повышало вероятность получения более правдоподобных ПЭС человека. Возраст всех дикторов находился в пределах от 18 до 54 лет, так как это возрастная группа характеризуется схожими устойчивыми эмоциональными реакциями. С каждым из дикторов было записано 24-и речевых фрагмента в исследуемых эмоциональных состояниях, а также по 3-и речевых фрагмента в спокойном состоянии. В результате итоговый набор содержал в себе 324 записи.

Второй набор записей включал в себя реальные эмоции. Для этого набора были собраны эмоциональные фразы, взятые из различных телевизионных передач, снятых в реальной обстановке. При этом исходный видеосигнал был в цифровом формате, что обеспечивало высокое качество речевых фрагментов в стандарте MPEG. В результате были получены фразы от 10 дикторов (5 мужчин и 5 женщин), отличных от тех, что принимали участие в записи первого набора. Возраст дикторов в данном наборе находился в пределах от 22 до 62 лет, что отражалось непосредственно в видеоматериале. Количество выделенных эмоциональных фраз изменялось от диктора к диктору и составляло до 5-ти для каждой эмоции. Общее число фраз составило 136 реализаций.

Проведены исследования различных параметров РС с целью выявления ПЭС человека, как в частотной, так и во временной области, в частности были исследованы коэффициенты ЛП, коэффициенты импульсной характеристики прямого фильтра и параметры ФП.

Определены зависимости изменения основных гармонических составляющих спектра РС при различных ПЭС человека, таких как частота основного тона (ОТ) и параметры первых четырех формантных частот. Показано, что эти параметры не обеспечивают высокую точность классификации ПЭС человека.

Проведены исследования зависимостей параметров РС от ПЭС человека во временной области. В качестве исследуемых параметров выступали длительность озвученного участка, интенсивность РС, джиттер и шиммер. Показано, что длительность озвученного участка РС, а также интенсивность РС обладают существенными различиями и обладают достаточной устойчивостью при различных ПЭС человека в случае совместного использования с другими параметрами РС.

Показаны изменения мел-кепстральных коэффициентов (МКК) при различных ПЭС человека. Показано, что МКК являются неустойчивыми с точки зрения классификации ПЭС человека, что объясняется дикторозависимостью этих коэффициентов.

Предложено использование коэффициентов ЛП и связанных с ними характеристик для классификации ПЭС человека. В качестве связанных характеристик выступали: коэффициенты отражения, коэффициенты импульсной характеристики прямого фильтра, автокорреляционные коэффициенты обратного фильтра, коэффициенты площади, логарифмические коэффициенты площади. Показано, что коэффициенты импульсной характеристики прямого фильтра и коэффициенты ЛП на основе автокорреляционного метода обладают высокой устойчивостью с точки зрения классификации ПЭС человека.

Предложено использование параметров ФП для классификации ПЭС человека. В качестве параметров ФП выступали коэффициенты асимметрии и эксцесса для сигнала (ось абсцисс) и его производной (ось ординат). Показано, что коэффициент эксцесса по оси ординат существенно изменяется при различных ПЭС человека.

Проведены исследования динамических параметров РС с целью выявления основных ПЭС человека. Для анализа динамики спектра РС предложено использовать алгоритм, построенный на основе фонетической функции А. А. Пирогова, с помощью которого оценивалась динамика выбранных параметров РС. В общем виде фонетическая функция на интервале времени будет выглядеть следующим образом:

$$P(\omega, t) = \int_0^{T_{\max}} \exp(-\tau_1/T) \lg \frac{S(\omega, t)}{S(\omega, t - \tau_1)} d\tau_1, \quad (1)$$

где $S(\cdot)$ – спектральная плотность РС в моменты времени t и $t - \tau_1$ в диапазоне частот ω , T – постоянная времени слухового аппарата, T_{\max} – длительность анализируемого РС. Рекомендуемый интервал времени τ_1 , определяющий соседние спектральные разрезы, обычно составляет: $\tau_1 = 20$ мс. Для

дискретного сигнала формула (1) будет иметь следующий вид:

$$P_k = \sum_{i=1}^{N_k} \lg \frac{S_k(i)}{S_{k-1}(i)}, \quad (2)$$

где, N_k – число отсчетов в кадре длительностью τ_1 , k – номер кадра.

Аналогично выражению для фонетической функции можно оценить динамику других параметров РС в следующем виде:

$$\hat{P}_k = \sum_{i=1}^{N_k} \lg \frac{Q_k(i)}{Q_{k-1}(i)}, \quad (3)$$

где $Q(\bullet)$ – вектор исследуемого параметра РС.

Показано, что все исследуемые динамические параметры РС изменяются под действием эмоций. При этом поведение частоты ОТ отличается для мужских и для женских голосов, а также в начальном и конечном участке фразы, что затрудняет использование данного параметра при классификации ПЭС человека.

Произведена оценка значимости статических и динамических параметров РС при классификации ПЭС человека. Для анализа эффективности статических и динамических признаков в задачах классификации ПЭС человека предложено использование комбинированного критерия, основанного на фонетической функции А. А. Пирогова, имеющего следующий вид:

$$\hat{Z}_k(a) = \ln \left(g_k / [g_{k-1}]^a \right), \quad k = \overline{1, N-1}. \quad (4)$$

где $g(\bullet)$ – исследуемый параметр РС, $a = 0 \dots 1$.

Комбинированная функция позволяет оценить вклад статических и динамических признаков в результирующий показатель. При этом их соотношение зависит от коэффициента a , так при $a = 0$ на результат оказывает влияние только статические параметры, а при $a = 1$ – только динамические. Таким образом, с ростом весового коэффициента a увеличивается влияние динамического параметра на результирующее значение комбинированной функции. Для оценки отклонения исследуемого параметра при различных ПЭС от спокойного состояния использовалось расстояние Евклида. Показано, что в случае оценки ПЭС человека статические параметры РС обладают большей значимостью ($a < 0,5$), чем динамические и являются более эффективными для классификации ПЭС человека.

Произведена оценка значимости полученных параметров РС с использованием МГУА, который использует рекурсивный селективный отбор моделей, на основе которых строятся более сложные модели.

Для оценки значимости полученных параметров РС предлагается использовать МГУА в том случае, когда выборка содержит несколько элементов. Метод использует индуктивный подход, согласно которому последовательно порождаются модели возрастающей сложности до тех пор, пока не будет найден минимум некоторого критерия качества модели.

В общем виде модель использует полином Колмогорова-Габора:

$$y = A_0 + \sum_{i=1}^m A_i K_i + \sum_{i=1}^m \sum_{j=1}^m A_{ij} K_i K_j + \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^m A_{ijk} K_i K_j K_k + \dots, \quad (5)$$

где Em – целевая эмоция, K_1, K_2, \dots, K_m – параметры РС.

Моделирование проводилось по двум независимым шкалам «удовольствие-неудовольствие» и «расслабление-напряжение», представленным на эмоциональной плоскости (рисунок 1). При этом условные эмоции на шкалах были приведены к значениям от -2,5 до 2,5 единиц. В качестве исследуемого массива РС выступал набор спровоцированных эмоций.

В результате анализа с помощью МГУА установлено, что наибольшей значимостью обладают коэффициенты эксцесса по оси ординат, коэффициенты импульсной характеристики прямого фильтра, коэффициенты ЛП, рассчитанные с помощью автокорреляционного метода, средняя интенсивность РС, параметры ФП и длительность озвученного участка. При этом средняя интенсивность и длительность озвученного участка использовались в известных ранее работах, в то время как остальные значимые параметры РС не встречались в работах, посвященных классификации ПЭС человека.

Исходя из анализа представленных зависимостей, получены следующие модели:

1 Для шкалы «удовольствие-неудовольствие»

$$Y_1 = A_0 + A_1 \cdot K_{im5}^2 + A_2 \cdot K_{im6} + A_3 \cdot K_{im8} \cdot \chi_{II}, \quad (6)$$

где $K_{im i}$ – i -е номера коэффициентов импульсной характеристики прямого фильтра χ_{II} – значения коэффициента эксцесса по оси ординат, $A_0 = 0,0097$, $A_1 = -0,0816$, $A_2 = 0,3551$, $A_3 = 0,0217$.

2 Для шкалы «расслабление-напряжение»

$$Y_2 = A_0 + A_1 \cdot \bar{E}_{dB} + A_2 \cdot K_{ar13} \cdot \bar{E}_{dB} + A_3 \cdot t_s + A_4 \cdot t_s \cdot K_{ar13} + A_5 \cdot t_s^2 + A_6 \cdot K_{ar13} + A_7 \cdot K_{ar13} \cdot K_{im2} + A_8 \cdot K_{im2} \cdot K_{im3}, \quad (7)$$

где $K_{ar i}$ – i -е номера коэффициентов ЛП, рассчитанные с помощью автокорреляционного метода, t_s – длительность озвученного участка, $A_0 = -0,3197$, $A_1 = -0,0470$, $A_2 = 1,0857$, $A_3 = 7,3206$, $A_4 = 6,8380$, $A_5 = -3,3400$, $A_6 = -57,2758$, $A_7 = -19,7213$, $A_8 = 0,0691$.

Представленные модели позволяют классифицировать эмоции по шкале «удовольствие-неудовольствие» с вероятностью 0,93, а по шкале «расслабление-напряжение» – с вероятностью 0,83.

Во второй главе произведена оценка влияния различных мешающих факторов в КПИ на значимые параметры РС с целью выявления ПЭС человека.

Определена необходимая полоса частот РС, содержащая признаки ПЭС человека. Показано, что для каждой эмоции характерны свои полосы частот (Таблица 1), при этом наибольшая полоса частот, необходимая для передачи

всех ПЭС человека составляет 0,3...6,5 кГц.

Проведены исследования возможности оценки ПЭС человека при действии акустических помех. В качестве акустических помех выступали широкополосная акустическая помеха (ШПАП), узкополосная акустическая помеха (УПАП) и импульсная акустическая помеха (ИАП).

Таблица 1

Эмоция	Нижняя граница, Гц	Верхняя граница, кГц
Радость	350	6,5
Ликование	650	3,5
Боязнь	330	5,5
Страх	350	6,5
Возмущение	300	4,5
Гнев	450	4,5
Наслаждение	320	6,5
Блаженство	380	5,5

Показано, что в случае действия ШПАП с уменьшением отношения сигнал-помеха (ОСП), снижается качество эмоциональной составляющей РС, а также вероятность классификации ПЭС человека. В случае действия УПАП снижение вероятности правильной классификации с ростом уровня шумов в средней снижается на 0,1. В случае действия ИАП вероятность правильной классификации ПЭС человека в среднем снижается на 0,11.

Произведена оценка возможности использования представления Хургина-Яковлева для повышения восприятия передаваемых эмоциональных составляющих РС. Показано, что применение представления второго порядка повышает оценку качества эмоциональной составляющей РС в среднем на 0,5 балла согласно ГОСТ Р 50840-95, а вероятности классификации ПЭС человека в среднем на 0,04.

Проведены исследования влияния параметров КПИ на эмоциональную составляющую РС. В качестве исследуемых кодеков источника выступали широко используемые кодеки AMR-NB, AMR-WB (G.722.2), CELP (G.728), ADPCM (G.726), G.723.1, AAC, обеспечивающие скорости передачи от 4,47 до 64 кбит/с соответственно, а в качестве канальных кодеков – сверточные и блочные коды. Установлено, что наибольшее влияние на эмоциональную составляющую РС оказывают кодеки источника, которые вносят изменения в параметры РС. При этом требуется использовать модифицированные модели классификации ПЭС человека на основе анализа РС, рассчитанные для каждого кодека в отдельности. Вероятности классификации ПЭС человека в случае использования модифицированных моделей представлены в таблице 2.

Таблица 2

Кодек	Шкала «удовольствие-неудовольствие»	Шкала «расслабление-напряжение»
AMR-WB	0,82	0,91
ADPCM	0,85	0,83
AAC	0,86	0,84

Проведены исследования влияния алгоритмов AMR на эмоциональную составляющую РС. Результаты исследований показали, что известные алгоритмы AMR обладают высокими показателями защиты, но при этом значительно ухудшают оценку ПЭС человека.

Предложено использование алгоритма AMR на основе представления Хургина-Яковлева при сохранении эмоциональной составляющей РС. Показано, что

алгоритм Хургина-Яковлева позволяет осуществить эффективное маскирование РС с качеством защищенного РС не более 1 балла, при этом восстановленный РС характеризуется качеством сравнимым с качеством незащищенных систем и обеспечивает вероятность классификации ПЭС человека 0,93 и 0,83 в случае отсутствия помех для шкал «удовольствие-неудовольствие» и «расслабление-напряжение» соответственно.

В третьей главе произведено обоснование и анализ реализации алгоритма классификации ПЭС человека на основе анализа РС на выходе КПИ.

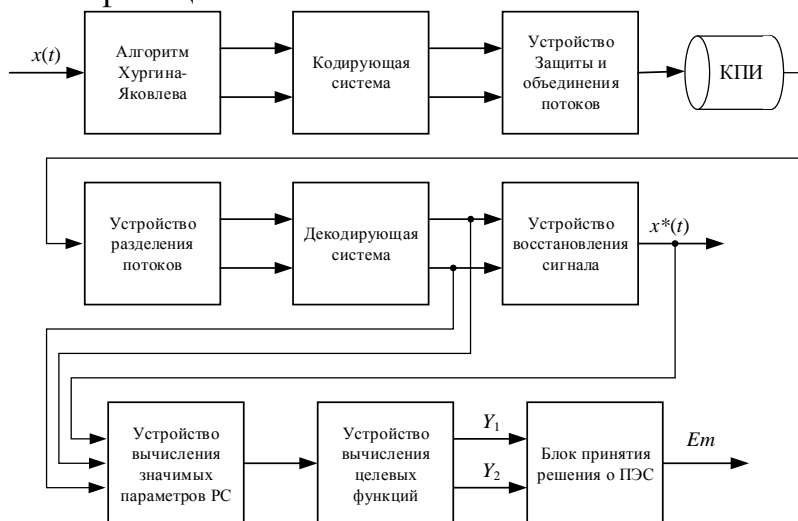


Рисунок 2

Обоснована структурная схема алгоритма классификации ПЭС человека, представленная на рисунке 2, состоящая из двух частей: алгоритма снижения влияния акустических помех и структурной защиты РС на основе представления Хургина-Яковлева и алгоритма оценки значимых параметров РС и

классификации ПЭС человека, блок схема которого представлена на рисунке 3.

Показана эффективность предложенного алгоритма маскирования, позволяющего обеспечить более высокую точность восстановления РС по сравнению с алгоритмом на основе теоремы В. А. Котельникова.

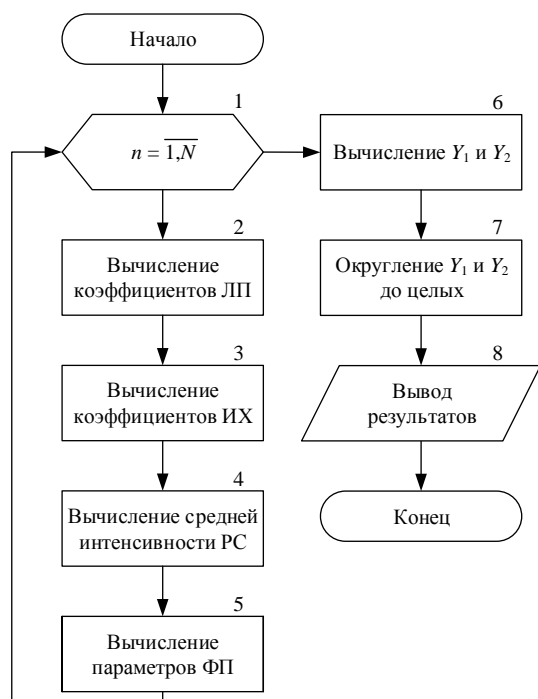


Рисунок 3

выигрыш при использовании представления Хургина-Яковлева составил 16 % при заданном числе уровней квантования и 13 % при исследовании величины коэффициента компрессии по сравнению с алгоритмом на основе теоремы В. А. Котельникова.

Показана эффективность предложенного алгоритма оценки значимых параметров РС и классификации ПЭС человека, позволяющего обеспечить вероятность точной классификации ПЭС человека равную 0,91 для шкалы «удовольствие-неудовольствие» и 0,80 для шкалы «расслабление-напряжение» при разрядности квантования более 16 и коэффициенте уменьшения разрядности более 0,7.

Произведен расчет вычислительной сложности аппаратной реализации алгоритма классификации ПЭС на выходе КПИ характеризуемой числом вычислительных операций, выполняемых за единицу времени и количеством используемых ячеек памяти. Показана возможность реализации алгоритма в режиме реального времени.

Произведен выбор цифрового сигнального процессора (ЦСП) по комбинированному критерию, учитывающему технические и экономические показатели качества, которому соответствовала микросхема TMS320C5535 фирмы Texas Instruments.

Предложено использование в качестве первичного кодека адаптированного ADPCM на процессорах семейства TMS320C55x, что позволяет эффективно использовать вычислительные возможности ЦСП как для задач кодирования, так и для реализации алгоритма классификации ПЭС на основе анализа РС на выходе КПИ. Предложенный кодек позволяет передавать РС с частотой дискретизации 16 кГц. Максимальные вычислительные затраты при этом не превышают 800 вычислительных операций в секунду.

ЗАКЛЮЧЕНИЕ ДИССЕРТАЦИОННОЙ РАБОТЫ

В заключении приведены основные результаты диссертационной работы.

1 Показано, что статические параметры РС имеют большой диапазон значений при различных ПЭС человека. При этом такие используемые ранее параметры РС, как значения частота ОТ и форманты не позволяют определить наличие ПЭС у человека с высокой вероятностью правильной классификации. Длительность озвученного участка и интенсивность РС позволяют оценить наличие эмоциональной окраски РС и являются достаточно устойчивыми при классификации ПЭС человека.

2 Проведено исследование динамических параметров РС, которое показало, что все исследуемые параметры изменяются под действием эмоций. При этом поведение частоты ОТ отличается для мужских и для женских голосов, а также в начальном и конечном участке фразы, что затрудняет использование данного параметра при классификации ПЭС человека.

3 Доказана более высокая значимость статических параметров РС по сравнению с динамическими при классификации ПЭС человека, с использованием комбинированной целевой функции на основе фонетической функции А.А. Пирогова.

4 Использование МГУА позволяет оценить значимость исследуемых параметров РС, в результате чего были построены модели классификации эмоций по двум шкалам, основанных на предложенной эмоциональной плоскости. При этом по шкале «удовольствие-неудовольствие» вероятность правильной классификации составила 0,93, а по шкале «расслабление-напряжение» – 0,83 при использовании спровоцированных эмоций, а при использовании реальных эмоций – 0,86 и 0,80 соответственно. Вероятность классификации ПЭС человека при схожем числе использованных ранее параметров составляла менее 0,80, а та же вероятность классификации

достигалась только при числе входных параметров более 15.

5 Установлено, что увеличение вероятности правильной классификации ПЭС человека обусловлено применением коэффициентов ЛП и связанных с ними параметров, а также использованием параметров ФП, построенных на основе использования представления Хургина-Яковлева.

6 Определена необходимая полоса частот КПИ для передачи РС, имеющего эмоциональный окрас, которая составила 0,3...6,5 кГц, что соответствует эффективной полосе передачи современных кодеков при частоте дискретизации 16 кГц.

7 Показано, что при действии акустических помех наибольшую опасность для автоматической классификации ПЭС человека представляют широкополосные и импульсные помехи. Показано, что с ростом уровня акустического шума возможно уменьшение вероятности правильной классификации ПЭС человека ниже 0,5.

8 Для повышения качества передаваемого РС и увеличения вероятности правильной классификации предложено использовать преобразование Хургина-Яковлева, в результате чего вероятность правильной классификации ПЭС человека при воздействии для ШПАП увеличивается в среднем на 0,05, УПАП в диапазоне 2000...3000 Гц – 0,03 и ИАП – 0,04.

9 Показано, что при передаче РС по КПИ наибольшее влияние оказывают кодеки источника, изменяющие параметры РС. При использовании адаптированных моделей классификации для каждого кодека вероятность классификации ПЭС человека повышается с 0,82 до 0,86 для эмоций группы «удовольствие-неудовольствие» и с 0,83 до 0,91 для эмоций группы «расслабление-напряжение».

10 Проанализировано влияние алгоритмов АМР на эмоциональную составляющую РС. Результаты показали, что современные алгоритмы АМР обладают высокими показателями защиты, но при этом не позволяют достаточно точно классифицировать ПЭС человека. Для сохранения эмоциональной составляющей РС предложено использовать алгоритм АМР на основе представления Хургина-Яковлева. Показано, что алгоритм Хургина-Яковлева позволяет осуществить эффективное маскирование РС с качеством защищенного РС не более 1 балла, при этом восстановленный РС по качеству сравнимым с РС в незащищенных системах и в случае отсутствия помех обеспечивает вероятность классификации ПЭС человека 0,93 и 0,83 для шкал «удовольствие-неудовольствие» и «расслабление-напряжение» соответственно.

11 Предложена система классификации ПЭС человека на основе анализа РС на выходе КПИ. Общая структурная схема алгоритма представлена в виде двух частей: алгоритма снижения влияния акустических шумов и защиты РС на основе представления Хургина-Яковлева и алгоритма классификации ПЭС человека. Предложенный алгоритм позволяет осуществить классификацию ПЭС человека на выходе КПИ с вероятностью до 0,93 для шкалы «удовольствие-неудовольствие» и 0,83 для шкалы «расслабление-напряжение».

12 Исследовано влияние шумов квантования в КПИ на алгоритм снижения влияния акустических шумов и защиты РС на основе представления Хургина-

Яковлева. Показана эффективность предложенного алгоритма, позволяющего обеспечить более высокую точность восстановления РС по сравнению с алгоритмом на основе теоремы В. А. Котельникова. Средний выигрыш при использовании представления Хургина-Яковлева составил 16 % при заданном числе уровней квантования и 13 % при исследовании величины коэффициента компрессии по сравнению с алгоритмом на основе теоремы В. А. Котельникова.

13 Исследовано влияние шумов квантования в КПИ на вероятность правильной классификации ПЭС человека на основе анализа РС. Показана эффективность предложенного алгоритма, позволяющего обеспечить вероятность точной классификации ПЭС человека равную 0,91 для шкалы «удовольствие-неудовольствие» и 0,80 для шкалы «расслабление-напряжение» при разрядности квантования более 16 и коэффициенте уменьшения разрядности более 0,7.

14 Проанализирована возможность аппаратной реализации алгоритма классификации ПЭС человека. Вычислительные затраты составили менее 900000 вычислительных операций в секунду и 620 слов памяти для хранения переменных и параметров. Рассмотрена возможность аппаратной реализации алгоритма классификации ПЭС человека на импортной и отечественной элементной базе. Анализ качественных характеристик ЦСП с помощью комбинированного критерия показал возможность построения алгоритма классификации ПЭС человека на выходе КПИ на микросхеме TMS320C5535 фирмы Texas Instruments. Показано возможность работы алгоритма в режиме реального времени.

15 Предложено использовать в качестве первичного кодека ADPCM на процессорах семейства TMS320C55x, что позволяет эффективно использовать вычислительные возможности ЦПС как для задач кодирования, так и для реализации алгоритма классификации ПЭС на основе анализа РС на выходе КПИ. Максимальные вычислительные затраты при этом не превышают 800 вычислительных операций в секунду.

Таким образом, разработанный алгоритм классификации ПЭС человека на основе анализа РС на выходе КПИ найдет применение в устройствах постобработки РС, а также средствах передачи и отображения медико-биологической информации.

СПИСОК ОПУБЛИКОВАННЫХ РАБОТ ПО ТЕМЕ ДИССЕРТАЦИИ

Публикации в журналах, рекомендованных ВАК

1. Дмитриев В.Т., Лукьянов Д.И. Алгоритм маскирования на основе представления Хургина-Яковлева с использованием производных второго и третьего порядков // Вестник РГРТУ. – 2012. – №42. – С. 13-17.
2. Кириллов С.Н. Лукьянов Д.И. Исследование влияния эмоционального состояния человека на параметры речевого сигнала // Вестник РГРТУ. – Рязань, 2014. – № 48. – С. 45-51.
3. Кириллов С.Н., Лукьянов Д.И. Исследование возможности оценки психоэмоционального состояния человека посредством анализа речевого

сигнала при действии акустических помех сигнала // Вестник РГРТУ. – Рязань, 2015. – № 51. – С. 19-23.

Публикации в межвузовских сборниках

4. Лукьянов Д.И. Оценка влияния характеристик систем передачи информации на речевой сигнал, имеющий эмоциональный окрас // Межвузовский сборник научных трудов «Методы и средства обработки и хранения информации». – Рязань, 2014. – С. 94-98.

Тезисы докладов на конференциях и семинарах

5. Дмитриев В.Т., Лукьянов Д.И. Использование представления Хургина-Яковлева в системах обработки речевых сигналов при использовании отсчетов сигнала и первых трех производных. // Семнадцатая международная научно-техническая конференция «Проблемы передачи и обработки информации в сетях и системах телекоммуникаций»: Тез. докладов. – Рязань: РГРТУ, 2012. – С. 67-68.

6. Дмитриев В.Т., Лукьянов Д.И. Применение представления Хургина-Яковлева в неинвазивной диагностике голосового тракта. // Двадцать четвертая всероссийская научно-техническая конференция студентов, молодых ученых и специалистов «Биотехнические, медицинские и экологические системы и комплексы»: Тез. докладов. – Рязань: РГРТУ, 2012. – С. 86-87.

7. Лукьянов Д.И. Алгоритм маскирования, выявления фальсификаций и восстановления фонограмм на основе представления Хургина-Яковлева. // Пятьдесят девятая студенческая научно-техническая конференция: Тез. докладов. – Рязань: РГРТУ, 2012. – С. 7-8.

8. Дмитриев В.Т., Лукьянов Д.И. Исследование алгоритма обработки речевых сигналов на основе представления Хургина-Яковлева для защиты фонограмм от фальсификаций. // Семнадцатая всероссийская научно-техническая конференция студентов, молодых ученых и специалистов: Тез. докладов. – Рязань: РГРТУ, 2012. – С. 65-67.

9. Лукьянов Д.И. Реализация алгоритма защиты, выявления фальсификаций и восстановления фонограмм на основе представления Хургина-Яковлева. // Вторая итоговая конференция молодых инноваторов по программе «Умник» в Рязанской области: Тез. докладов. – Рязань: РГРТУ, 2012. – С. 21-23.

10. Лукьянов Д.И. Распознавание психоэмоционального состояния человека на основе параметров речи. // Двадцать пятая всероссийская научно-техническая конференция студентов, молодых ученых и специалистов «Биотехнические, медицинские и экологические системы и комплексы»: Тез. докладов. – Рязань: РГРТУ, 2012. – С. 192-193.

11. Дмитриев В.Т., Семин Д.С., Виноградова М.Е., Лукьянов Д.И., Картавенко Я.О. Определение информативных параметров речи для выявления эмоционального состояния диктора и качества речи при влиянии различных акустических помех // Десятая международная научная конференция «Перспективные технологии в средствах передачи информации»: Тез. докладов. – Владимир: ПТСПИ, 2013. – С. 151-152.

12. Лукьянов Д.И. Оценка интенсивности эмоционального состояния космонавтов посредством речевого сообщения. // Шестая международная научно-техническая конференция «Космонавтика. Радиоэлектроника. Геоинформатика»: Тез. докладов. – Рязань: РГРТУ, 2013. – С. 168-169.

13. Лукьянов Д.И. Определение значимых параметров речевого сообщения для выявления психоэмоционального состояния человека по шкале удовольствие-неудовольствие // Восемнадцатая всероссийская научно-техническая конференция студентов, молодых ученых и специалистов «Новые информационные технологии в научных исследованиях»: Тез. докладов. – Рязань: РГРТУ, 2013. – С. 103-104.

14. Кириллов С.Н., Покровский П.С., Семин Д.С., Лукьянов Д.И., Яшин А.Ю., Лисничук А.А. Адаптивная к мешающим факторам, помехозащищенная радиопередача информации // Шестнадцатая международная конференция «Цифровая обработка сигналов и ее применение – DSPA-2014»: Тез. докладов. – Москва, 2014. – С. 80-84.

15. Дмитриев В.Т., Лукьянов Д.И. Использование представления Хургина-Яковлева для создания закрытого формата записи и хранения речевых сообщений // Семнадцатая международная телекоммуникационная конференция молодых ученых и студентов «Молодежь и наука». Тезисы докладов. – Москва, 2014. – С. 33.

16. Акопов Э.В., Лисничук А.А., Лукьянов Д.И., Яшин А.Ю. Адаптивные системы передачи информационных сообщений, устойчивая к воздействию мешающих факторов // Международная конференция «Радиоэлектронные устройства и системы для инфокоммуникационных технологий – РЭУС-2014». Тезисы докладов. – Москва, 2014. – С. 186-190.

17. Лукьянов Д.И., Лисничук А.А. Анализ влияния структуры систем передачи речевой информации на качество оценки психоэмоционального состояния человека // Четвертая всероссийская научно-техническая конференция «Информационно-измерительные и управляющие системы военной техники». Тезисы докладов. – Владимир, 2014. – С. 100-102.

18. Лукьянов Д.И. Оценка коэффициентов линейного предсказания речевого сигнала в зависимости от психоэмоционального состояния человека // Девятнадцатая всероссийская научно-техническая конференция студентов, молодых ученых и специалистов «Новые информационные технологии в научных исследованиях»: Тез. докладов. – Рязань: РГРТУ, 2014. – С. 115-116.

19. Лукьянов Д.И. Система оценки психоэмоционального состояния пациента при удаленном общении. // Двадцать седьмая всероссийская научно-техническая конференция студентов, молодых ученых и специалистов «Биотехнические, медицинские и экологические системы и комплексы»: Тез. докладов. – Рязань: РГРТУ, 2014. – С. 38-39.

20. Лукьянов Д.И. Оценка возможности использования фазовых портретов для определения психоэмоционального состояния диктора // Восемнадцатая международная телекоммуникационная конференция молодых ученых и студентов «Молодежь и наука». Тезисы докладов. – Москва, 2015. – С. 22-23.

Лукьянов Дмитрий Игоревич

**Алгоритмы обработки речевых сигналов и классификации
психоэмоционального состояния человека на выходе каналов
передачи речевой информации при действии мешающих
факторов**

Автореферат

диссертации на соискание ученой степени

кандидата технических наук

Подписано в печать 20.04.2015. Формат бумаги 60x84 1/16.

Бумага офсетная. Печать трафаретная. Усл. печ. л. 1,0.

Тираж 100 экз. Заказ №

Отпечатано в НПЦ «Информационные технологии».

390035, г. Рязань, ул. Островского, д. 21/1.