

Маркин Дмитрий Николаевич

РАЗРАБОТКА МЕТОДА И ТЕХНИЧЕСКИХ СРЕДСТВ
КОМПАНДИРОВАНИЯ СПЕКТРОВ РЕЧЕВЫХ СИГНАЛОВ

Специальность 05.11.18

Приборы и методы преобразования
изображений и звука

Автореферат диссертации
на соискание ученой степени
кандидата технических наук

Санкт-Петербург – 2008

Работа выполнена на кафедре акустики Санкт-Петербургского
государственного университета кино и телевидения

Научный руководитель – доктор технических наук, профессор Ува-
ров Владимир Константинович

Официальные оппоненты:

доктор технических наук Майоров Василий Семенович;

кандидат технических наук Плющёв Владимир Михайлович

Ведущая организация – ОАО НПП «Дальняя связь»

Защита состоится «23» декабря 2008 г. в ____ часов на заседании дис-
сертационного совета Д 210.021.01 в Санкт-Петербургском государствен-
ном университете кино и телевидения по адресу:

191119, Санкт-Петербург, ул. Правды, д. 13

С диссертацией можно ознакомиться в библиотеке Санкт-Петербургского
государственного университета кино и телевидения.

Автореферат разослан «__» ноября 2008 г.

Ученый секретарь диссертационного совета
кандидат технических наук, профессор

К.Ф. Гласман

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность исследования.

Число средств передачи информации по радиоканалу непрерывно возрастает. При этом речевая коммуникация (телефония) до сих пор остается одним из наиболее оперативных и востребованных способов непосредственного обмена информацией между людьми. На сегодняшний день число абонентов средств радиотелефонии значительно превышает количество абонентов фиксированных сетей общего пользования, поэтому проблема эффективного использования радиочастотного спектра становится все острее, так как радиочастотный спектр является ограниченным невозобновляемым ресурсом. Одним из путей эффективного использования радиочастотного ресурса является сжатие спектра передаваемых сигналов, в частности речевых, как занимающих значительную долю среди всех транслируемых по радиоканалу сигналов.

Несмотря на то, что проблема компандирования спектра речевых сигналов (РС) на сегодняшний день достаточно успешно решается средствами статистической теории с устранением психоакустической избыточности, поиск решений данной проблемы на базе альтернативных теоретических представлений не только не потерял своей актуальности, но и приобрел еще большую остроту с развитием телекоммуникационных технологий, что объясняется ограниченными возможностями известных методов при возрастающей потребности. Одним из теоретических фундаментов, который традиционно рассматривается в качестве основы для решения обозначенной проблемы, является модуляционная теория звуковых сигналов, в этом направлении ориентирована данная работа.

Разработка новых эффективных способов компандирования спектра РС является актуальной, прежде всего, для систем радиосвязи, в том числе специализированных систем подвижной радиосвязи. Также это актуально для систем записи и хранения больших массивов речевой информации.

Объект исследования: обработка речевых сигналов.

Предмет исследования: Компандирование (сжатие – восстановление) спектра речевых сигналов при их обработке на основе математической модели модуляционной теории.

Основная цель исследования заключается в разработке эффективного метода компандирования спектра речевых сигналов.

В соответствии с основной целью и предметом исследования определены следующие основные **задачи исследования**:

- экспериментально исследовать свойства функций, описывающих речевые сигналы в математической модели модуляционной теории;
- разработать новую психоакустическую модель, учитывающую особенности слухового восприятия модулированных по амплитуде и частоте сигналов;
- разработать метод и технические средства компандирования спектра речевых сигналов при обработке их природных модулирующих функций;
- экспериментально исследовать эффективность разработанного метода компандирования спектра речевых сигналов;
- обеспечить внедрение результатов, полученных в диссертационном исследовании, в учебный процесс и практику научных исследований.

Методологическую и теоретическую основы исследования составили научные труды отечественных и зарубежных авторов в теории функций, физиологии слуха и речи, теории информации, спектральной теории и модуляционной теории.

Методы исследования. Во время проведения исследования применялись методы математического анализа; спектрального анализа (теоретического и экспериментального); компьютерного моделирования; артикуляционных измерений, и др.

Научная новизна исследования:

1. На основе экспериментальных измерений параметров мгновенной частоты речевых сигналов внесены уточнения в сделанные ранее выводы о

свойствах мгновенной частоты фоном русского языка: во-первых, модулирующие частоты природной частотной модуляции гласных звуков русской речи могут принимать значения выше 1 кГц, во-вторых, природная несущая частота гласных звуков русской речи, как правило, не равняется частоте основного тона голоса.

2. Для устранения психофизической избыточности речевых сигналов с целью сжатия их спектра предложена новая психоакустическая модель, учитывающая особенности слухового восприятия модулированных по амплитуде и частоте сигналов.

3. Для русского языка экспериментально установлено, что модулирующие частоты природных амплитудной и частотной модуляций речевого сигнала в значительной степени совпадают.

4. Для русского языка экспериментально установлено, что при сохранении информации о величине природной несущей частоты речевого сигнала можно незаметно для слуха подменить составляющие спектра сигнала, обусловленные его природной частотной модуляцией, аналогичными спектральными составляющими, обусловленными его природной амплитудной модуляцией.

5. На основе предложенной психоакустической модели разработан способ и устройство компандирования спектра речевого сигнала при делении мгновенных частот сигнала и его огибающей и ограничении сверху спектров огибающей огибающей, косинусов фазы сигнала и огибающей с деленными мгновенными частотами. Экспериментально оценена эффективность компандирования спектра речи разработанным способом.

Научная ценность результатов исследования заключается в том, что полученные в ходе работы при выполнении данного диссертационного исследования научные и технические результаты (например, предложенная новая психоакустическая модель слухового восприятия, установленные факты совпадения и возможности незаметной для слуха подмены модулирующих частот природных амплитудной и частотной модуляций речевых

сигналов) могут быть использованы при создании новых методов и средств обработки речевых сигналов на основе математической модели модуляционной теории.

Практическая значимость работы. Проведенные теоретические и экспериментальные исследования позволили разработать эффективный способ и устройство компандирования спектров РС при обработке их природных модулирующих функций. Реализация цифровыми средствами разработанного способа компандирования спектра РС в современных системах радиосвязи позволит увеличить эффективность использования радиочастотного ресурса.

Реализация результатов диссертационной работы. Материалы диссертационной работы внедрены в учебный процесс и научно-исследовательскую работу Института ФСБ России (г. Санкт-Петербург) и кафедры акустики Санкт-Петербургского государственного университета кино и телевидения, а также в НИР ООО «Неватон». Внедрение результатов диссертационной работы подтверждено соответствующими актами.

Апробация работы. Основные положения диссертационной работы докладывались и обсуждались в трех сообщениях на научно-технических конференциях Санкт-Петербургского государственного университета кино и телевидения в 2006-2007 годах.

Публикации. По материалам выполненных исследований опубликовано 11 печатных работ, в которых изложено основное содержание диссертации: 10 статей, и одна заявка на «Способ и устройство сжатия спектра речевых сигналов», по которой получен патент РФ на изобретение.

Личный вклад автора. Все основные научные результаты, изложенные в диссертационной работе, получены автором лично. Все публикации написаны автором диссертации под редакцией и при общем научном руководстве соавторов.

Структура и объем работы. Диссертация содержит введение, основной текст из четырех глав, заключение, список использованной литературы

и приложение. Объем основного текста с введением и заключением составляет 193 страницы, включая 57 рисунков на 46 страницах и 2 таблицы на 2 страницах. Список использованной литературы содержит 138 наименований.

На защиту выносятся следующие положения:

1. Для сжатия спектра речевого сигнала на основе математической модели модуляционной теории необходимо сжимать спектры огибающей и косинуса фазы сигнала.
2. Модулирующие частоты природной амплитудной и природной частотной модуляций речевых сигналов в значительной степени совпадают.
3. Установлена возможность в определенных пределах незаметной для человеческого слуха подмены модулирующих частот природной частотной модуляции на аналогичные модулирующие частоты природной амплитудной модуляции речевых сигналов.

КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении обоснована актуальность темы диссертации, сформулированы проблема, основная цель и задачи исследования, дана характеристика работы, сформулированы основные положения, выносимые на защиту, описана структура диссертации.

В первой главе с позиций теории информации, теории спектров и модуляционной теории звуковых сигналов обозначена проблема сжатия спектра РС. Обоснована необходимость продолжения развития модуляционной теории звуковых сигналов, изучающей свойства натуральных акустических сигналов. Обоснована необходимость сжатия спектра речевых сигналов для повышения эффективности использования частотного ресурса каналов передачи речи. Показано развитие и современное состояние решения проблемы компандирования спектра РС с целью их трансляции по каналам связи. Приведены зависимости качества речи от степени компрессии спектра РС наиболее популярными современными методами.

Сжатие спектра РС возможно за счет уменьшения их статистической и

психоакустической избыточностей. В современных системах радиотелефонии с целью сжатия спектра речевых сигналов наиболее широкое применение нашли гибридные вокодеры, уменьшающие как психоакустическую, так и статистическую избыточности. Достаточно низкое качество получаемой речи при сравнительно невысокой степени сжатия ее спектра современными методами обосновывает необходимость поиска новых путей эффективного решения данной проблемы на базе альтернативных теоретических представлений.

Проведен обзор исследовательских работ, посвященных решению проблемы сжатия спектра РС на базе модуляционных представлений. На основе анализа приведенных работ в конце первой главы делаются выводы о возможности решения проблемы компандирования спектра РС на основе модуляционной теории и формулируются основные задачи настоящего диссертационного исследования: исследовать свойства функций, описывающих РС в математической модели модуляционной теории; разработать методы сжатия спектров огибающей и косинуса фазы РС.

Во второй главе приводятся результаты экспериментальных исследований спектральных и информационных свойств огибающей и косинуса фазы речевых сигналов.

Экспериментально исследованы соотношения между спектром речевого сигнала и спектрами субсигналов его огибающей и косинуса фазы, выделяемых в канале неполного модуляционного анализа. Установлено, что верхние граничные частоты спектров огибающей и косинуса фазы РС практически соответствуют верхней граничной частоте спектра самого сигнала. Также установлено, что значительная часть информации, содержащейся в речевом сигнале, одновременно отражена в его огибающей, и в косинусе фазы, то есть дублируется (рис. 1), а часть информации о речевом сигнале отражена либо только в его огибающей, либо только в косинусе фазы. Информация о значениях природных модулирующих частот речевого сигнала в высокой степени дублируется (модулирующие частоты природ-

ных амплитудной и частотной модуляций в значительной мере совпадают). Информация о значении несущей частоты речевого сигнала и о ее относительно медленном изменении при переходе от одного звука речи к другому содержится только в косинусе фазы (мгновенной частоте).

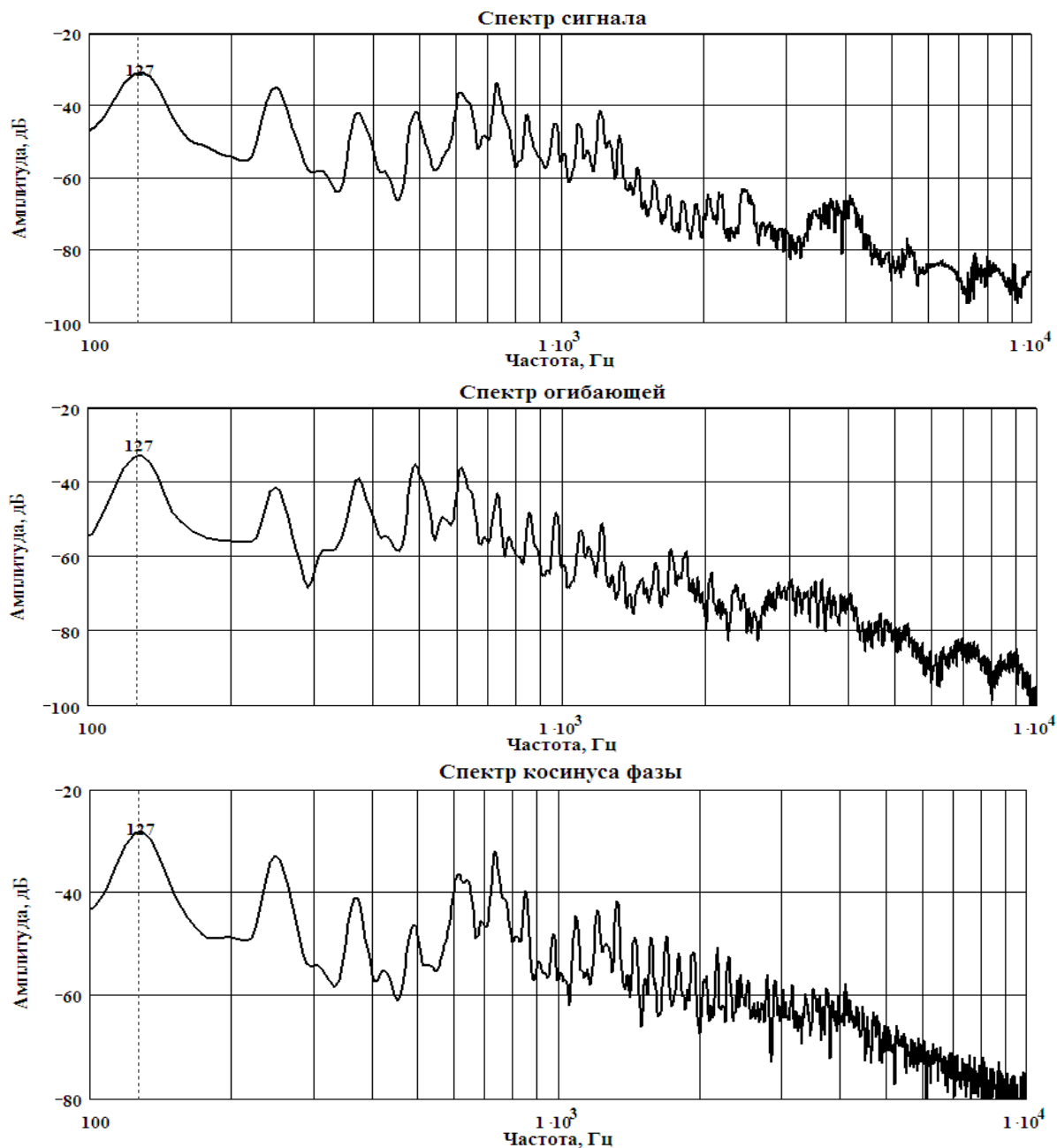
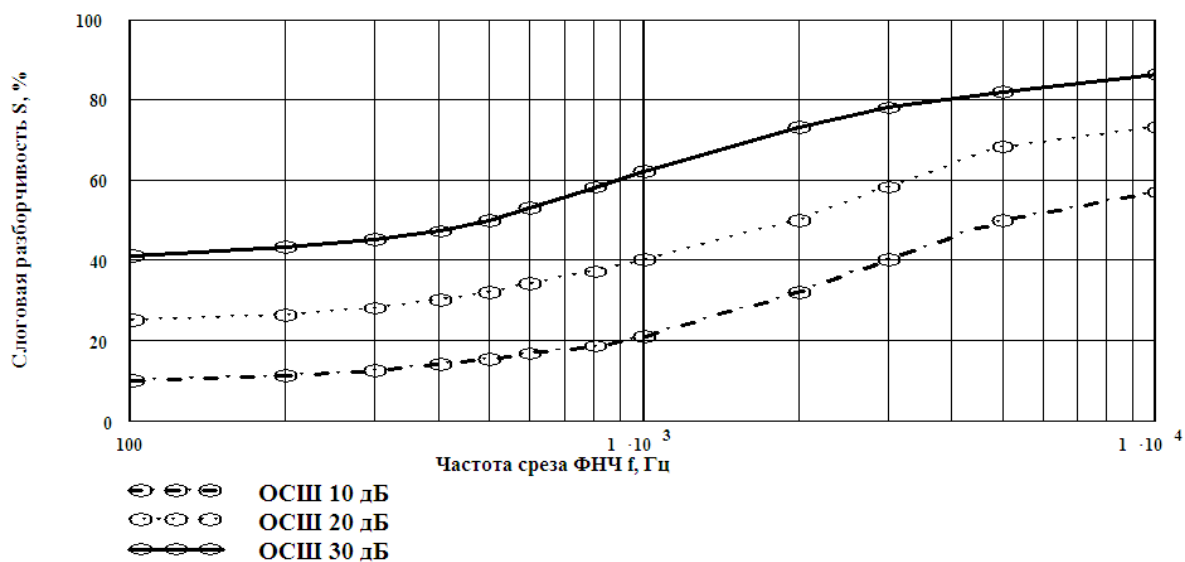


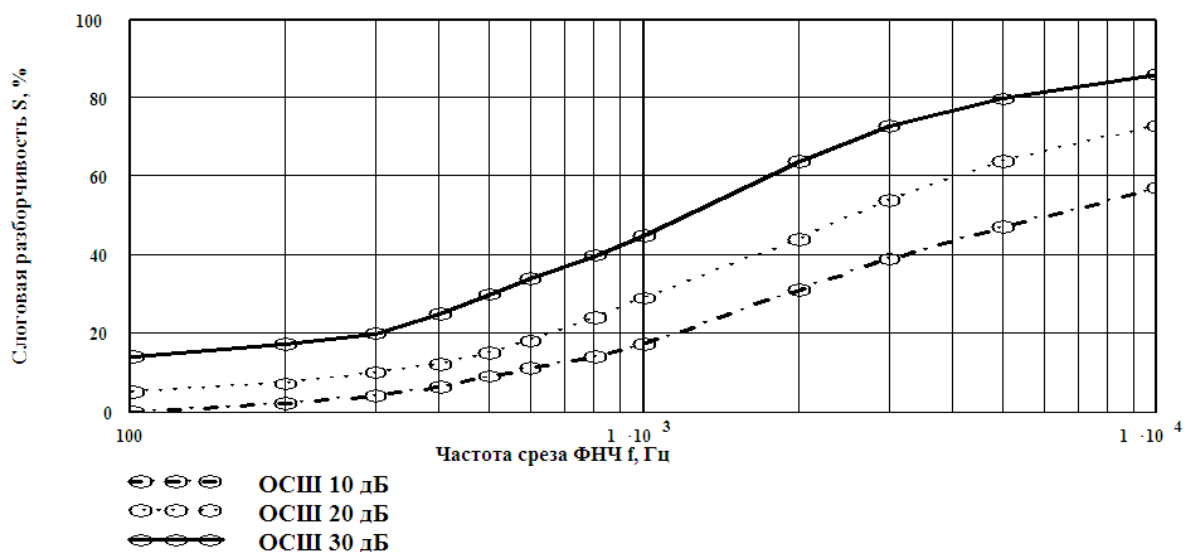
Рис. 1. Спектры речевого сигнала, его огибающей и косинуса фазы для звука «А»

Обоснован выбор определения разборчивости речи по слоговой артикуляции. Измерены зависимости разборчивости речи от верхних граничных частот спектров косинуса фазы и огибающей РС (рис. 2, 3) при их раздельном ограничении сверху с помощью ФНЧ. Все эксперименты по определению слоговой разборчивости проводились в строгом соответствии с

требованиями ГОСТ Р 50840 – 95. По полученным результатам делается вывод, что косинус фазы речевого сигнала играет бóльшую роль в передаче семантической информации речи, так как разборчивость речи быстрее снижается при уменьшении верхней граничной частоты спектра косинуса фазы, то есть высокочастотные компоненты спектра косинуса фазы более информативны, чем аналогичные компоненты спектра огибающей РС. При ограничении сверху спектра косинуса фазы сигнала теряется не только информация о модулирующих частотах природной частотной модуляции и девиации мгновенной частоты РС, но и информация о величине природной несущей частоты сигнала, что существенно влияет на разборчивость речи.



ис. 2. Зависимость слоговой разборчивости речи от верхней граничной частоты спектра огибающей речевого сигнала при исходном косинусе фазы (для разных отношений сигнал/шум – ОСШ)



ис. 3. Зависимость слоговой разборчивости речи от верхней граничной частоты спектра

косинуса фазы речевого сигнала при исходной огибающей (для разных ОСШ)

Экспериментально установлено, что временные сдвиги между субсигналами огибающей и косинуса фазы речевых сигналов при их отдельной обработке в канале неполного модуляционного анализа-синтеза оказывают заметное влияние на качество и разборчивость синтезируемой репродукции речи, так как такие относительные временные расхождения приводят к появлению частотных искажений в синтезированном сигнале.

По результатам экспериментальных исследований спектральных и информационных свойств огибающей и косинуса фазы РС делаются выводы о значении этих функций в передаче речевой информации. Косинус фазы играет большую роль в передаче информации о согласных фонемах, а огибающая – о гласных фонемах русского языка. Огибающая речевого сигнала больше отвечает за качество передачи речи. Косинус фазы речевого сигнала в большей мере определяет передачу семантической информации. Делается вывод о нецелесообразности применения ограничения сверху спектров косинуса фазы и огибающей РС для сокращения полосы частот канала связи, требующейся для трансляции речи. Делается вывод, что для сжатия спектра РС на основе модуляционной теории звуковых сигналов и передачи речи с коммерческим качеством (не ниже первого класса качества речи по ГОСТ Р 50840 – 95) необходимо передавать огибающую и косинус фазы сигнала и сжимать спектры как огибающей, так и косинуса фазы.

В третьей главе разрабатывается математическая основа способа сжатия спектра РС на базе математической модели модуляционной теории.

Прежде всего, математически строго обосновывается возможность применения преобразования Гильберта к звуковым сигналам (являющимся относительно широкополосными), на использовании которого строится модуляционная теория. Тем самым показывается, что для описания звуковых сигналов непротиворечиво и теоретически обоснованно может быть использована математическая модель модуляционной теории.

В математической модели модуляционной теории любой звуковой сиг-

нал $s(t)$ описывается в виде сложно-модулированного (одновременно по амплитуде и частоте) процесса, то есть в виде произведения огибающей $S(t)$ (амплитудно-модулирующей функции) и косинуса фазы $\cos\varphi(t)$ (частотно-модулированной функции) сигнала

$$s(t) = S(t) \cos \varphi(t) = S(t) \cos \int_0^t \omega(t) dt, \quad (1)$$

где $\varphi(t)$ – текущая фаза сигнала, $\omega(t)$ – мгновенная частота (частотно-модулирующая функция) сигнала. Функции $S(t)$ и $\cos\varphi(t)$ являются функциями первой ступени модуляционного разложения (неполного модуляционного анализа) сигнала.

Огибающая сигнала $s(t)$ определяется следующим образом

$$S(t) = \sqrt{s^2(t) + \hat{s}^2(t)}, \quad (2)$$

где $s(t)$ – исходный сигнал, $\hat{s}(t)$ – опорный сигнал, сопряженный по Гильберту с сигналом $s(t)$

$$\hat{s}(t) = H[s(t)] = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{s(\tau)}{t - \tau} d\tau. \quad (3)$$

Ранее* было экспериментально установлено, что выделение мгновенной частоты $\omega(t)$ звуковых сигналов в виде субсигнала даже без выполнения с ней каких-либо преобразований приводит к неустраняемым искажениям в синтезируемой на ее основе репродукции звукового сигнала. Поэтому для обработки мгновенной частоты звуковых сигналов отказались от необходимости ее детектирования, были разработаны методы и средства для обработки модулирующих функций $S(t)$ и $\omega(t)$ звуковых сигналов без их детектирования. Также было установлено, что без детектирования частотно-модулирующей функции $\omega(t)$ сигнала невозможно сжать ее спектр, так как она стоит аргументом под функцией косинуса.

В работе для сжатия спектра огибающей и косинуса фазы речевого сигнала предлагается углубить разложение сигнала по модулирующим функциям, описывать сигнал в рамках двух ступеней модуляционного разложения. Из формулы (2) видно, что огибающая сигнала является неотри-

* Уваров В.К. Точное компандирование частотного и динамического диапазонов звуковых сигналов. – СПб.: СПбГУКиТ, 2001. – 326 с.

цательной функцией времени, в ее спектре содержится ненулевая постоянная составляющая. В соответствии со спектральными представлениями спектр огибающей можно разделить на две смежные области – низкочастотную и высокочастотную, поэтому в работе функцию $S(t)$ предлагается представлять в виде суммы неотрицательной низкочастотной и знакопеременной высокочастотной частей (рис. 4)

$$S(t) = S_{\text{нч}}(t) + S_{\text{вч}}(t). \quad (4)$$

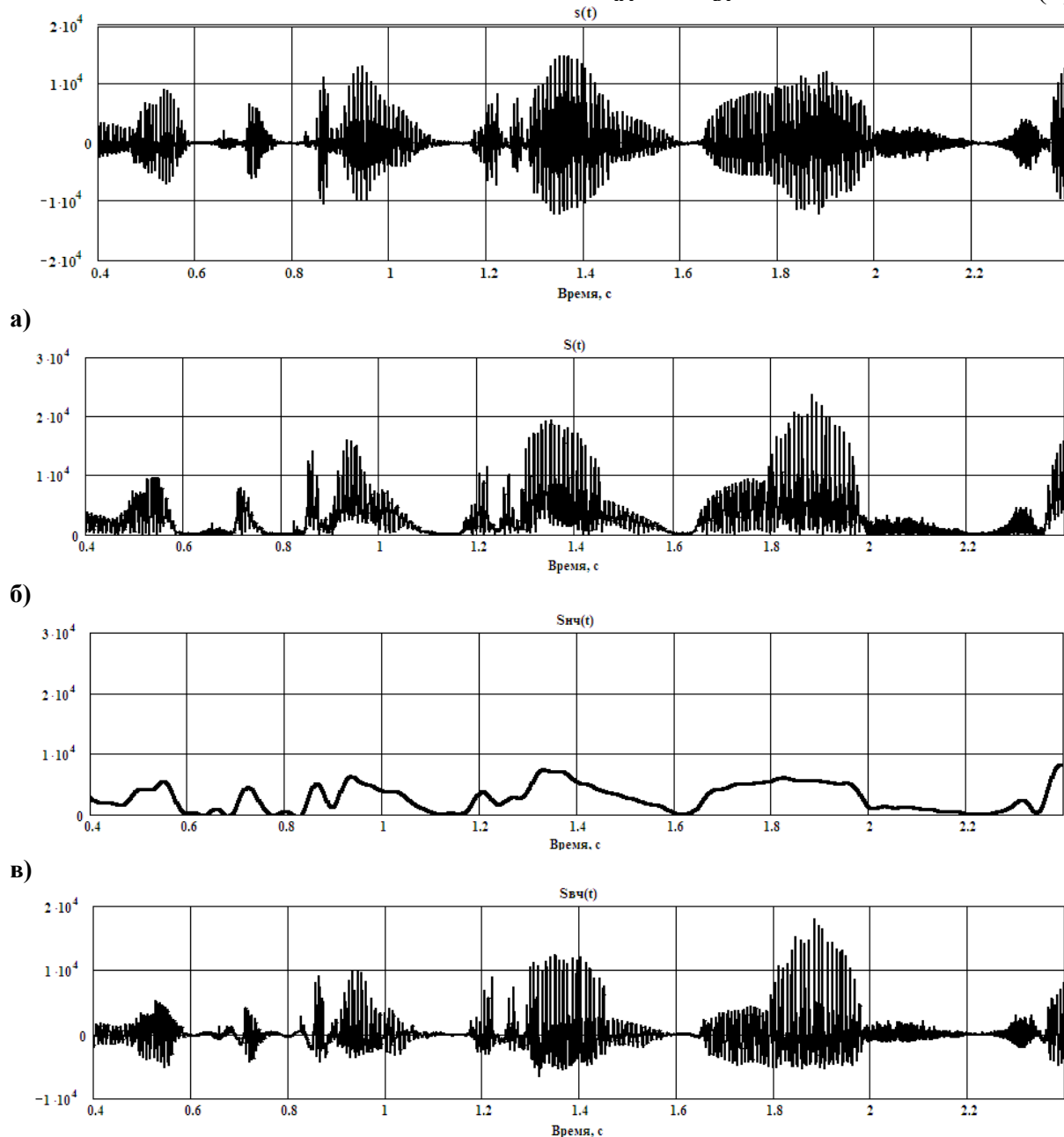


Рис. 4. Разделение огибающей речевого сигнала на низкочастотную и высокочастотную составляющие: а) речевого сигнала; б) огибающая речевого сигнала; в) низкочастотная составляющая огибающей сигнала; г) высокочастотная составляющая огибающей сигнала

В соответствии с модуляционными представлениями высокочастотная

составляющая огибающей может быть представлена в виде

$$S_{вч}(t) = S_s(t) \cos \varphi_s(t), \quad (5)$$

где $S_s(t)$ и $\cos \varphi_s(t)$ – огибающая и косинус фазы огибающей сигнала, функции второй ступени модуляционного разложения сигнала. Следовательно огибающую речевого сигнала можно записать в виде

$$S(t) = S_{нч}(t) + S_{вч}(t) = S_{нч}(t) + S_s(t) \cos \varphi_s(t). \quad (6)$$

В работе предложено речевые сигналы описывать функциями двух ступеней модуляционного разложения следующим образом –

$$s(t) = [S_{нч}(t) + S_s(t) \cos \varphi_s(t)] \cos \int_0^t \left(\omega_0(t) + \omega_d(t) \cos \int_0^t \omega_m(t) dt \right) dt, \quad (7)$$

где $\omega_0(t)$ – природная несущая частота сигнала, $\omega_d(t)$ – девиация мгновенной частоты сигнала, $\omega_m(t)$ – модулирующая частота природной частотной модуляции сигнала. Описание речевого сигнала в виде сложно-модулированного процесса моделью (7) совпадает по виду с формулой при объединении классических амплитудной и частотной модуляций, при модуляции гармоническими сигналами

$$s(t) = [U_0 + U_m \cos \omega_{ам} t] \cos \int_0^t (\omega_0 + \omega_d \cos \omega_{чм} t) dt. \quad (8)$$

Разница в том, что в формуле (7) все параметры сигнала являются функциями времени, а в модели (8) – постоянными величинами.

Разработан способ и устройство для двухступенного модуляционного разложения речевых сигналов в соответствии с их представлением моделью (7). Структурная схема разработанного способа представлена на рис. 5. Для двухступенного модуляционного разложения РС предлагается последовательно выполнить два раза процедуру неполного модуляционного анализа сигнала. Приведены характеристики и описание работы отдельных блоков и узлов разработанного устройства для выполнения двух ступеней модуляционного разложения речевых сигналов.

В четвертой главе, прежде всего, экспериментально определены условия выполнения второй ступени модуляционного разложения РС: 1) оп-

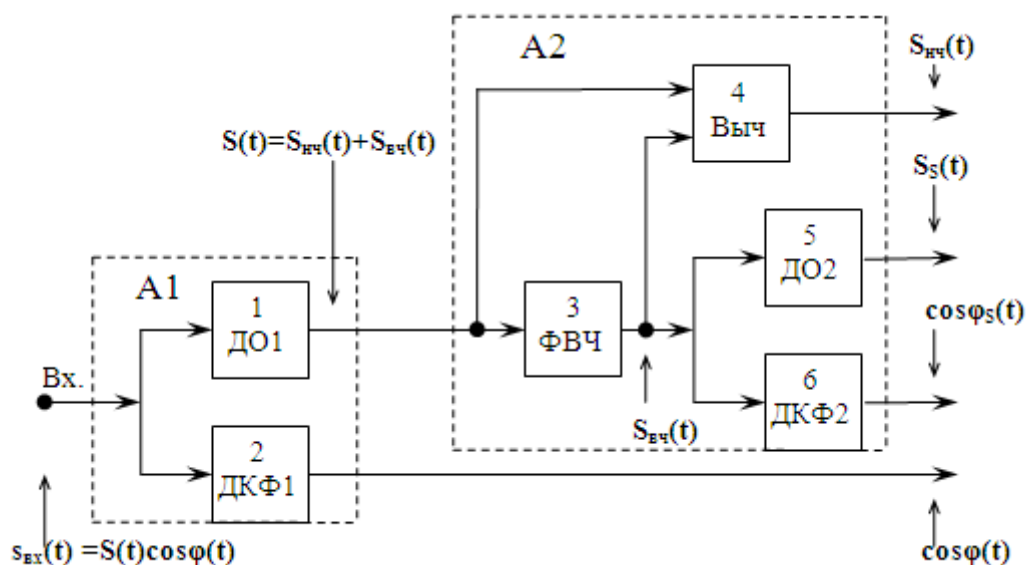


Рис. 5. Структурная схема устройства для двухступенного модуляционного разложения речевых сигналов. $A1$, $A2$ – первый и второй анализаторы; $ДО1$, $ДО2$ – детекторы огибающей; $ДКФ1$, $ДКФ2$ – детекторы косинуса фазы; $ФВЧ$ – фильтр высоких частот с граничной частотой 15 Гц; $Выч$ – вычитатель сигналов

ределены оптимальные с точки зрения обеспечения качества передачи речи параметры и схема устройства для двухступенного модуляционного разложения РС; 2) экспериментально установлено, что преобразование Гильберта огибающей РС необходимо выполнять в полосе частот приблизительно 15 Гц – 4 кГц (при ограничении спектра речевого сигнала верхней граничной частотой 4 кГц); 3) установлено, что для обеспечения коммерческого качества передачи речи (не ниже первого класса качества по ГОСТ Р 50840 – 95) при двухступенном модуляционном разложении необходимо передавать низкочастотную часть огибающей РС $S_{нч}(t)$.

Экспериментально исследованы спектральные и информационные свойства огибающей, косинуса фазы и мгновенной частоты огибающей РС. Установлено, что верхние границы спектров огибающей и косинуса фазы огибающей близки к верхней границе спектра самого сигнала. Также установлено, что модулирующие частоты природных частотных модуляций сигнала и его огибающей в значительной степени совпадают. Часть информации дублируется в модулирующих функциях второй ступени модуляционного разложения речевого сигнала.

Для некоторых фонем русского языка, для мужского и женского голо-

сов проведено измерение основных параметров (при описании моделью частотной модуляции) мгновенной частоты речевого сигнала и мгновенной частоты огибающей сигнала. Исследование мгновенной частоты огибающей РС выполнено впервые. Экспериментально установлено, что: во-первых, модулирующие частоты природной частотной модуляции гласных фонем русского языка могут принимать значения превышающие 1 кГц, во-вторых, для всех фонем русского языка природная несущая частота сигнала превышает несущую частоту огибающей сигнала, в-третьих, несущие частоты звуков русской речи могут принимать значения приблизительно от ста герц до нескольких килогерц, а модулирующие частоты – приблизительно от семидесяти герц до нескольких килогерц, причём низким модулирующим частотам соответствуют низкие значения несущей частоты.

Теоретически и экспериментально исследовано влияние ограничения сверху спектра косинуса фазы сигнала без деления мгновенной частоты и с делением мгновенной частоты. Установлено, что при соответствующем подборе коэффициента деления мгновенной частоты и верхней граничной частоты фильтра низких частот ограничение сверху спектра косинуса фазы с деленной мгновенной частотой после восстановления масштаба мгновенной частоты эквивалентно ограничению сверху спектра мгновенной частоты, при этом сохраняется величина природной несущей частоты и часть модулирующих частот природной частотной модуляции сигнала.

В работе для сжатия спектра РС на базе математической модели модуляционной теории при устранении психофизической избыточности сигнала предложена новая психоакустическая модель, учитывающая особенности слухового восприятия модулированных по амплитуде и частоте сигналов. Ранее* было экспериментально установлено, что человеческий слух различает амплитудную и частотную модуляции только в пределах одной частотной группы (в случае равенства несущих и модулирующих частот амплитудной и частотной модуляций, когда энергетические спектры частот-

*Цвикер Э., Фельдкеллер Р. Ухо как приёмник информации. – М.: Связь, 1971. – 255 с.

но- и амплитудно- модулированного сигналов совпадают). На рис. 6 представлена экспериментально полученная зависимость слухового различения амплитудной и частотной модуляций от значения несущей частоты.

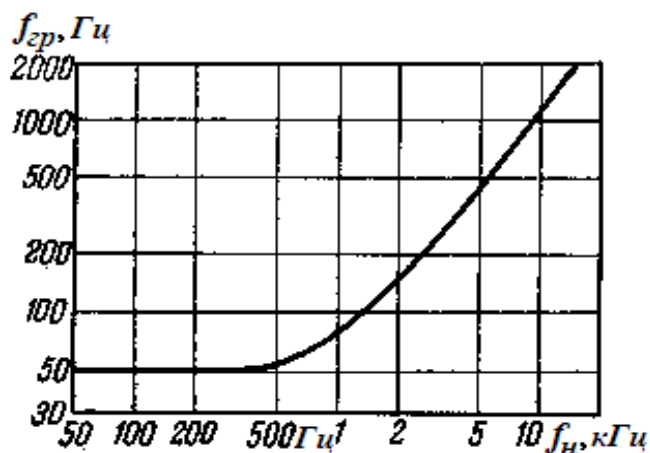


Рис. 6. Зависимость граничной частоты слуховой дифференциации амплитудной и частотной модуляций от несущей частоты (по [Цвикер Э., Фельдкеллер Р. Ухо как приёмник информации. – М.: Связь, 1971, с. 190])

В области спектральных частот, где сосредоточена основная часть энергии речевых сигналов (100 – 3000 Гц) ширина критической полосы чувствительности к разнице амплитудной и частотной модуляций составляет приблизительно 50 – 250 Гц (рис. 6), для большей части значений модулирующих частот природных модуляций речевых сигналов спектральные продукты амплитудной и частотной модуляций выходят за пределы одной частотной группы, то есть не различаются на слух.

Экспериментально установлено, что при сохранении информации о величине природной несущей частоты незаметно для человеческого слуха можно подменить природную частотную модуляцию речевых сигналов их природной амплитудной модуляцией, так как их модулирующие частоты в значительной степени совпадают, а спектральные продукты выходят за пределы одной частотной группы.

На основе экспериментально полученных данных о свойствах функций двух ступеней модуляционного разложения РС и предложенной психоакустической модели разработан способ и устройство компандирования спектра РС (структурная схема которого изображена на рис. 7) при делении мгновенной частоты сигнала и его огибающей и ограничении сверху спек-

тров огибающей огибающей, косинусов фазы сигнала и огибающей с деленными мгновенными частотами. При такой обработке сигнала частично ограничивается сверху спектр частотно-модулирующей функции сигнала, сохраняется информация о величине природной несущей частоты РС и ее изменении при переходе от одного звука речи к другому. Информация о природных модулирующих частотах речевого сигнала передается, в основном, в амплитудно-модулирующей функции РС, чей спектр сжимается на передающей стороне и восстанавливается на приемной стороне. На приемной стороне часть модулирующих частот природной частотной модуляции незаметно для слуха подменяется модулирующими частотами природной амплитудной модуляции сигнала. На разработанные способ и устройство получен патент РФ на изобретение.

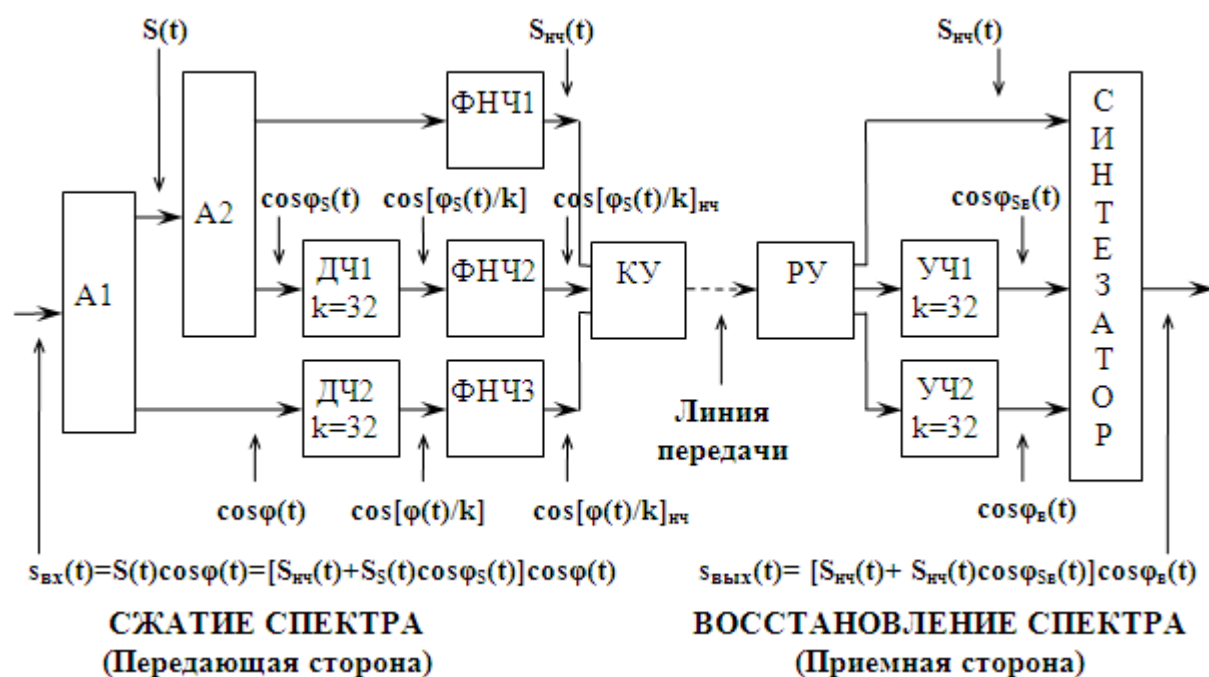


Рис. 7. Структурная схема способа компрессирования спектра речевых сигналов при обработке модулирующих функций двух ступеней модуляционного разложения. А1, А2 – первый и второй анализаторы (рис. 5); ДЧ1, ДЧ2 – делители мгновенной частоты в 32 раза; ФНЧ1, ФНЧ2, ФНЧ3 – фильтры низких частот; КУ – комбинирующее устройство; РУ – распределительное устройство; УЧ1, УЧ2 – умножители мгновенной частоты в 32 раза; СИНТЕЗАТОР – синтезатор, состоящий из двух перемножителей и сумматора

На передающей стороне сигнал разделяется на функции: $S_{\text{нч}}(t)$, $\cos \varphi_s(t)$, $\cos \varphi(t)$. После этого мгновенная частота сигнала и его огибающей после делится в 32 раза, спектры косинусов фазы сигнала и его огибающей после

деления мгновенной частоты ограничиваются сверху с помощью фильтров низких частот, полученные таким образом функции после их объединения методами вторичного уплотнения в комбинирующем устройстве передаются по каналу связи. На приемной стороне восстанавливается масштаб мгновенных частот сигнала и его огибающей, после чего по формуле

$$s_{\text{вых}}(t) = [S_{\text{нч}}(t) + S_{\text{нч}}(t)\cos\varphi_{\text{св}}(t)]\cos\varphi_{\text{в}}(t) \quad (9)$$

формируется репродукция речевого сигнала.

На рис. 8 приведена измеренная зависимость слоговой разборчивости речи от полосы частот канала связи при компандировании спектра РС разработанным способом (2). По этим результатам оценена эффективность разработанного способа компандирования спектра речевых сигналов.

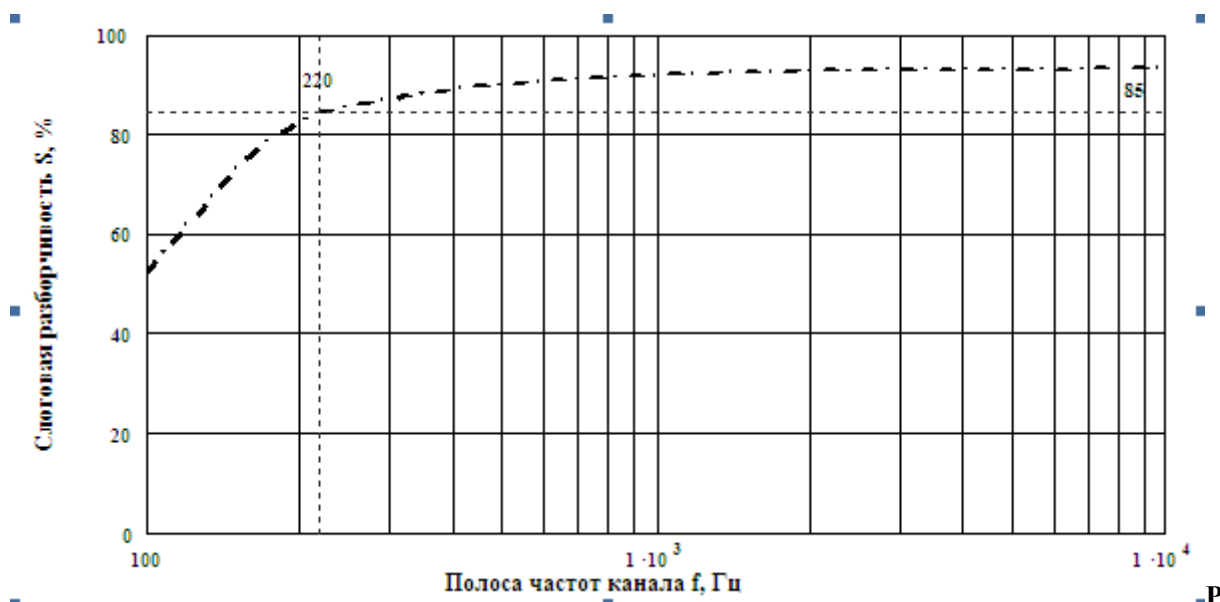


рис. 8. Зависимость слоговой разборчивости речи от полосы частот канала связи при сжатии спектра речевого сигнала разработанным способом

Сжимая спектр РС разработанным способом, полосу частот канала связи можно сократить примерно до 220 Гц, слоговая разборчивость при этом составит не ниже 86%, что соответствует первому классу качества речи по разборчивости в соответствии с ГОСТ Р 50840 – 95 для систем передачи речи с пониженными скоростями.

ЗАКЛЮЧЕНИЕ

1. На основе результатов экспериментальных исследований сделаны выводы, что для обеспечения коммерческого качества передачи речи (не

ниже первого класса качества в соответствии с ГОСТ Р 50840 – 95) и решения задачи эффективного компандирования спектра речевых сигналов на основе модуляционной теории необходимо передавать огибающую и косинус фазы, и сжимать спектры огибающей, и косинуса фазы сигнала.

2. Экспериментально установлено, что модулирующие частоты природной амплитудной и природной частотной модуляций речевых сигналов в значительной степени совпадают.

3. Для сжатия спектра речевых сигналов на основе модуляционной теории при устранении психофизической избыточности сигнала предложена новая психоакустическая модель, учитывающая особенности слухового восприятия модулированных по амплитуде и частоте процессов. Человеческий слух различает амплитудную и частотную модуляции только в пределах одной частотной группы слуха.

4. Экспериментально установлено, что при сохранении информации о величине природной несущей частоты речевых сигналов можно незаметно для слуха подменить модулирующие частоты природной частотной модуляции на аналогичные модулирующие частоты природной амплитудной модуляции речевых сигналов.

5. На основе предложенной психоакустической модели разработан способ и устройство компандирования спектра речевых сигналов при обработке их модулирующих функций двух ступеней модуляционного разложения: при ограничении сверху спектра природной частотно-модулирующей функции сигнала, сохранении информации о величине природной несущей частоты сигнала и сжатии спектра природной амплитудно-модулирующей функции сигнала. Информация о природной модуляции речевого сигнала передается, в основном, в его природной амплитудно-модулирующей функции. На приемной стороне незаметно для слуха недостающие модулирующие частоты природной частотной модуляции речевого сигнала заменяются аналогичными модулирующими частотами природной амплитудной модуляции сигнала. На разработанные способ и устройство для его реализа-

ции получен патент РФ на изобретение.

6. Разработанный способ компандирования спектра речевого сигнала позволяет сократить полосу частот, требуемую для передачи речи по каналу связи с качеством не ниже первого класса по разборчивости в соответствии с ГОСТ Р 50840 – 95 (для систем передачи речи с пониженными скоростями), примерно до 220 Гц. Выигрыш в полосе частот канала связи при этом составляет около 15 раз (по отношению к полосе 3400 Гц, принятой для телефонных сетей общего пользования). При менее жестких требованиях к качеству речи и более высоких требованиях к степени компрессии (например, в частных коммерческих и специализированных системах связи) полосу частот канала связи можно сократить еще больше, например, приблизительно до 170 Гц, то есть в 20 раз, обеспечивая при этом второй класс качества по разборчивости речи в соответствии с ГОСТ Р 50840 – 95 (для систем передачи речи с пониженными скоростями).

7. Результаты диссертационного исследования внедрены в учебный процесс и научно-исследовательскую работу Института ФСБ России (г. Санкт-Петербург) и кафедры акустики Санкт-Петербургского государственного университета кино и телевидения, а также в НИР ООО «Нева-тон».

В диссертации поставлена и решена актуальная научно-техническая задача в области преобразования речевых сигналов. На основании выполненных исследований в диссертации изложены научно обоснованные технические разработки, имеющие значение для экономики страны. Разработаны способ и устройство компандирования спектра речевых сигналов при обработке их модулирующих функций двух ступеней модуляционного разложения.

ОСНОВНОЕ СОДЕРЖАНИЕ ДИССЕРТАЦИИ ОПУБЛИКОВАНО В СЛЕДУЮЩИХ РАБОТАХ

1. Маркин Д.Н., Уваров В.К. К вопросу о необходимой точности настройки звеньев фазовращателя в канале модуляционного анализа-синтеза//

Проблемы развития кинематографа и телевидения/Сборник научных трудов СПбГУКиТ, Вып. 18. – СПб.: Изд. СПбГУКиТ, 2005. С. 69 – 79.

2. Маркин Д.Н., Уваров В.К. Результаты новых исследований на основе модуляционной теории звуковых сигналов/В сб. Факультету аудиовизуальной техники – 75 лет. – СПб.: Изд. СПбГУКиТ, 2005. С. 36 – 41.

3. Маркин Д.Н., Уваров В.К. О применении преобразования Гильберта к звуковым сигналам. Деп. рук. № 186кт-Д07, ОНТИ НИКФИ, 2007. – 12 с.

4. Маркин Д.Н., Уваров В.К. Актуальность исследования свойств модулирующих функций огибающей речевых сигналов. Деп. рук. № 183кт-Д07, ОНТИ НИКФИ, 2007. – 22 с.

5. Маркин Д.Н., Уваров В.К. Результаты исследований, подтверждающих возможность технической реализации второй ступени модуляционного разложения речевых сигналов. Деп. рук. № 185кт-Д07, ОНТИ НИКФИ, 2007. – 26 с.

6. Маркин Д.Н., Уваров В.К. Математическая модель огибающей речевых сигналов и способ реализации второй ступени модуляционного разложения речевого сигнала. Деп. рук. № 187кт-Д07, ОНТИ НИКФИ, 2007. – 14 с.

7. Маркин Д.Н., Уваров В.К. Обзор исследований, посвящённых решению задачи сжатия спектра речевых сигналов на основе модуляционных представлений. Деп. рук. № 182кт-Д07, ОНТИ НИКФИ, 2007. – 24 с.

8. Маркин Д.Н., Уваров В.К. Результаты практических исследований соотношений между спектрами сигнала, его огибающей, косинуса фазы и мгновенной частоты. Деп. рук. № 181кт-Д07, ОНТИ НИКФИ, 2007. – 32 с.

9. Маркин Д.Н., Уваров В.К. Результаты практических исследований информационных свойств огибающей и косинуса фазы речевых сигналов. Деп. рук. № 188кт-Д07, ОНТИ НИКФИ, 2007. – 28 с.

10. Маркин Д.Н., Осташевский Е.Н., Уваров В.К. Время-частотный анализ сигналов. Деп. рук. № 184кт-Д07, ОНТИ НИКФИ, 2007. – 38 с.

11. Маркин Д.Н., Уваров В.К. Способ и устройство сжатия спектра речевых сигналов. Патент РФ на изобретение № 2320028. Приоритет от

27.09.2006. // БИ. № 8, 20.03.2008. – 11 с.