

Министерство образования Республики Беларусь
Учреждение образования
Белорусский государственный университет
информатики и радиоэлектроники

На правах рукописи

УДК 62-503.55

Коминч
Николай Юрьевич

**Повышение эффективности модели речевого кодера на основе
алгоритма Малла**

Автореферат
диссертации на соискание степени
магистра техники и технологии

по специальности 1-59 81 01 – Управление
безопасностью производственных процессов

Минск 2019

Работа выполнена на кафедре инженерной психологии и эргономики учреждения образования «Белорусский государственный университет информатики и радиоэлектроники»

Научный руководитель: **Камлач Павел Викторович,**
кандидат технических наук, доцент, доцент
кафедры экологии учреждения образования
«Белорусский государственный университет
информатики и радиоэлектроники»

Рецензент: **Майсеня Люмила Иосифовна,**
Кандидат физико-математических наук, доктор
педагогических наук, профессор, заведующая
кафедрой физико-математических дисциплин
института информационных технологий
«Белорусского государственного университета
информатики и радиоэлектроники»

Защита диссертации состоится «28» июня 2019 г. в 9.00 часов на заседании Государственной экзаменационной комиссии по защите магистерских диссертаций в учреждении образования «Белорусский государственный университет информатики и радиоэлектроники» по адресу:

С диссертацией можно ознакомиться в библиотеке учреждения образования «Белорусский государственный университет информатики и радиоэлектроники»

ВВЕДЕНИЕ

За последние десятилетия быстро развиваются методы обработки и анализа информационных сигналов с применением вейвлет-преобразования. Они позволяют исследовать свойства нестационарных сигналов, благодаря операциям масштабирования и сдвига. Разработаны вейвлеты для обработки речевых сигналов, обеспечивающие сжатие, декомпозицию, реставрацию, идентификацию речи, а также удаление из них шумов.

На данный момент задача качественной передачи оцифрованной речи со скоростью от 2 до 6 Кбит/с должным образом все еще не решена. Это обуславливает поиск новых алгоритмов эффективной низкоскоростной компрессии и декомпрессии речевых сигналов в реальном масштабе времени. Решение проблемы компрессии речевого сигнала особенно актуально в таких областях, как Интернет-телефония, сотовая связь, запись и хранение речевых сообщений в специальных и портативных устройствах.

Кодеры формы сигнала и схемы, основанные на линейном предсказании, используют временное представление сигнала и обеспечивают хорошее качество при скоростях передачи 32 Кбит/с и выше. В диапазоне скоростей ниже 16 Кбит/с, который требуется для беспроводных и спутниковых систем коммуникаций, при сохранении разборчивости у таких кодеров отмечается существенная деградация качества синтезируемой речи. При низких скоростях передачи информации более эффективным является параметрическое кодирование, в котором в качестве параметров используют различные характеристики представления сигнала в частотной области – спектральные кодеры [1], в частности использующие Вейвлетное преобразование.

Целью исследования является разработка метода повышения эффективности модели речевого кодера на основе алгоритма Малла при обработке, сжатии и восстановлении радиотехнических сигналов. Для достижения поставленной цели необходимо разработать алгоритм работы кодера, выбрать средства реализации проекта, реализовать схемы декомпозиции и реконструкции сигнала в программной среде, произвести моделирование процесса кодирования-декодирования речевого сигнала, разработать наиболее эффективный способ повышения качества кодирования речевого сигнала.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования.

Интенсивное развитие информационных технологий в современном мире делает особенно важным решение проблемы быстрой и качественной передачи различного рода информации по цифровым линиям связи. Несмотря на разнообразие применяемых для этого средств связи, основным видом коммуникации между людьми остается передача информации посредством речи. В связи с этим в настоящее время продолжают интенсивно развиваться и совершенствоваться методы цифровой обработки и передачи речевых сообщений. Использование цифрового представления данных позволяет обеспечить надежность и экономичность связи, возможность гарантированной защиты от несанкционированного доступа. Благодаря появлению новых средств обработки сигналов и соответствующих специальных инструментальных сред появились возможности быстрого создания перспективных систем передачи и обработки речевых сообщений. При этом большое значение приобретает решение проблемы минимизации числа бит необходимых для передачи сигнала, т.е. компрессии и кодирования речи. Одной из наиболее актуальных задач на сегодняшний день является создание систем низкоскоростной передачи с высоким качеством восприятия сигнала, способных корректно функционировать в реальных условиях окружающей среды.

Степень разработанности проблемы

На данный момент задача качественной передачи оцифрованной речи со скоростью от 2 до 6 Кбит/с должным образом все еще не решена. Это обуславливает поиск новых алгоритмов эффективной низкоскоростной компрессии и декомпрессии речевых сигналов в реальном масштабе времени. Решение проблемы компрессии речевого сигнала особенно актуально в таких областях, как Интернет-телефония, сотовая связь, запись и хранение речевых сообщений в специальных и портативных устройствах.

Цель и задачи исследования

Целью исследования является разработка методов повышения эффективности модели речевого кодера на основе алгоритма Малла при обработке, сжатии и восстановления радиотехнических сигналов.

Для достижения поставленной цели необходимо решить следующие задачи:

1. Разработать алгоритм работы кодера;
2. Выбрать средства реализации проекта;

3. Реализовать схемы декомпозиции и реконструкции сигнала в программной среде и произвести моделирование процесса кодирования-декодирования речевого сигнала;

4. Разработать наиболее эффективный способ повышения качества кодирования речевого сигнала.

Объектом исследования являются метод кодирования речевого сигнала на основе алгоритма Малла.

Предметом работы являются характеристики речевого сигнала и их изменение при кодировании и декодировании.

Область исследования. Основным объектом исследований является один из методов компрессии, основанный на применении к речевому сигналу дискретного Вейвлет-преобразования. Для этого в проекте будет разработан речевой кодер для быстрого Вейвлет-преобразования основанного на кратном анализе (алгоритм Малла) и исследованы его характеристики. Алгоритм будет реализован на основе ПЛИС в среде Quartus.

Теоретическая и методологическая основа исследования

В основу диссертации легли работы белорусских и зарубежных ученых в области кодирования речевых сигналов.

Информационная база исследования сформирована на основе литературы, открытой информации, технических нормативно-правовых актов, сведений из электронных ресурсов, а также материалов научных конференций и семинаров.

Теоретической основой исследований, проведенных в работе, являются общенаучные методы сравнительного анализа, методы оценки количественной и качественной эффективности кодирования речевых сигналов.

Методологической основой исследования являются разработки отечественных и зарубежных производителей систем кодирования, технической и иной документации на исследуемые системы. В магистерской диссертации используются следующие общенаучные методы: структурный и сравнительный анализ, метод формализации. В диссертации используется системный подход к разработке архитектуры комплексной системы компрессии.

Инструментальной базой исследований являются пакеты прикладных программ Quartus и Matlab.

Научная новизна и значимость полученных результатов обоснована необходимостью совершенствования методов и алгоритмов низкоскоростной компрессии речевых сигналов в реальном времени и целесообразностью разработки устройств для низкоскоростной компрессии речи. В свою очередь, для проектирования качественных низкоскоростных вокодерных

устройств необходимо использование не только новых алгоритмов сжатия речи, но и адаптированных для этих целей аппаратных средств.

Основные положения, выносимые на защиту

1. Анализ структуры VOIP.
2. Создание модели речевого кодера на основе алгоритма Малла.
3. Методы повышения эффективности Вейвлет преобразования.

Теоретическая значимость диссертации заключается в том, что в ней предложены несколько подходов для повышения эффективности компрессии речевого сигнала на основе алгоритмов Малла.

Практическая значимость диссертации состоит в том, что на основе предложенных методов возможно построить систему компрессии речевых сигналов с наиболее эффективными параметрами.

Структура и объем работы

Работа состоит из введения, трёх глав и заключения, библиографического списка. Общий объём диссертации – 100 страниц. Работа содержит 42 рисунка. Библиографический список включает 39 наименования.

В первой главе были рассмотрены основы организации IP-телефонии, пример архитектуры сети, а также описаны факторы влияющие на качество речи. Приведены основные патенты в данной области и сделан анализ актуальности темы магистерской диссертации.

Во второй главе была обоснована и смоделирована структурная схема речевого кодера и более детально рассмотрены его основные узлы. Была описана реализация КИХ-фильтра с составлением блок-схем регистров. Также представлены структурные схемы блоков декомпозиции и реконструкции сигнала.

В третьей главе в результате теоретических и экспериментальных исследований, а также математического и компьютерного моделирования описаны методы повышения эффективности вейвлет-преобразований при обработке, сжатии и восстановления радиотехнических сигналов.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении обоснована актуальность и практическая ценность темы исследования, описаны решаемые проблемы и цели исследования. Приведен обзор работ на данную тему, указано текущее состояние технологий кодирования речевых сигналов.

В общей характеристике работы показана актуальность проводимых исследований, степень разработанности проблемы, сформулированы цель и задачи диссертации, обозначена область исследований, научная (теоретическая и практическая) значимость исследований, а также апробация работы.

В первой главе рассмотрены основы организации IP-телефонии, пример архитектуры сети, а также описаны факторы влияющие на качество речи. Приведены основные патенты в данной области и сделан анализ актуальности темы магистерской диссертации.

Технология передачи речевой информации по сетям с маршрутизацией пакетов IP – Voice over IP (VOIP) или IP-телефонии представляет собой компрессию голосового сигнала с последующей передачей по цифровым каналам передачи данных с использованием протокола IP. Частным случаем IP-телефонии является Интернет-телефония, при которой в качестве передающей сети используется сеть Интернет.

Одно из основных требований, предъявляемых к сетям VOIP – способность подключаться к традиционным коммутируемым телефонным сетям общего пользования (ТфОП). Международный союз электросвязи (ITU) решил эту проблему, создав H.323 – комплект стандартов для мультимедийных сетей пакетной коммутации.

Непрерывное увеличение передач аудио и речевых данных в системах мультимедиа через Интернет обуславливает поиск новых решений цифровой обработки в реальном масштабе времени аудио и речевых сигналов (их компрессию и декомпрессию). Использование цифрового представления данных позволяет обеспечить надежность и экономичность связи, возможность гарантированной защиты от несанкционированного доступа. Благодаря появлению новых процессоров обработки сигналов появилась возможность быстрого создания перспективных систем передачи или запоминания речевых сообщений. При этом большое значение приобретает проблема минимизации числа бит, необходимых для передачи сигнала, проблема кодирования и компрессии речи. Актуальной задачей обработки речи стало создание систем низкоскоростной передачи с высоким качеством восприятия сигнала, способных функционировать в реальных условиях. В

системах речевой связи данные передаются, хранятся и обрабатываются различными способами, что обуславливает применение разных форм представления сигнала.

К ним предъявляются следующие требования:

- сохранение информационного содержания речи;
- удобство передачи и хранения;
- возможность легкого и гибкого преобразования без существенных информационных потерь (задача кодирования и декодирования речевого сигнала).

Целью цифрового кодирования речи является получение как можно более высокого качества восстанавливаемой речи при наименьшей скорости передачи. Исследования в области кодирования речи проводятся более трех десятков лет. Результат этих интенсивных усилий – множество различных стратегий и подходов для кодирования речевых сигналов, ряд которых доведен до соответствующих национальных и международных стандартов.

Во второй главе был спроектирован один из вариантов речевого кодера. Эффективный алгоритм реализации дискретного Вейвлет-преобразования был предложен Малла. Этот алгоритм является классической схемой, посредством которой входной одномерный сигнал раскладывается на высоко и низкочастотные компоненты с использованием пары ортогональных фильтров. Весь алгоритм основан на четырех действиях, а именно фильтрация, децимация, интерполяция и суммирование. Функционально фильтры представляют собой НЧ и ВЧ КИХ-фильтры, дециматор представляет собой устройство, которое понижает частоту дискретизации в 2 раза, интерполятор повышает частоту дискретизации в 2 раза. Алгоритм при этом обеспечивает трехуровневую компрессию речевого сигнала.

Точность реконструкции сигналов зависит от потерь информации при выполнении прореживания спектров, причем эти потери наблюдаются на срезах полос пропускания фильтров низких и высоких частот, крутизна которых зависит от порядка фильтров, их согласованности, и типа Вейвлетных функций. Обязательным условием преобразования сигнала является его задание количеством точек (отсчетов), равном $N=2m$, где значение $m \geq 1$ определяет максимально возможное число уровней декомпозиции сигнала при целочисленных значениях кратности сдвигов операторов фильтров количеству отсчетов Вейвлетных коэффициентов на каждом уровне декомпозиции. Для выполнения этого условия количество отсчетов сигнала, как правило, дополняется до ближайшего большего значения N методами, известными из практики задания начальных/конечных условий свертки

(нулями, концевыми значениями сигналов, четными или нечетными значениями относительно концевых отсчетов, периодическим продолжением и т.п.). Может применяться также передискретизация исходного сигнала до необходимого количества отсчетов.

Моделирование работы системы кодер-декодер производилось на основе программы Matlab использованием реальных речевых сигналов, ранее записанных в файл.

Результаты моделирования показали, что ошибки достигают 0,04 по амплитуде, а относительно исходных сигналов $0,04/0,9 = 0,044 = 4 \%$. Эти ошибки нормальны, так как происходило преобразование коэффициентов фильтра к целым двоичным числам до 2^3 , а также при реконструкции происходило деление, отбрасывались младшие разряды.

В третьей главе было предложено несколько способов повышения эффективности рассмотренного кодера на основе алгоритма Малла. Преобразование с помощью вейвлет-пакетов является адаптивным вейвлет-преобразованием, поскольку оно легко приспосабливается к особенностям сигнала и может успешно использоваться для компрессии сигналов и их очистки от шумов.

Достоинством вейвлет-пакетов и адаптивных алгоритмов их реализации является отсутствие необходимости в обучении системы и даже в оценке статистических характеристик сигналов. Все, что нужно - это ввести оценку стоимости вейвлет-коэффициентов, мерой которой может служить энтропия – концентрация числа вейвлет-коэффициентов M , требующихся для описания сигнала с некоторой заданной точностью (или погрешностью).

Для осуществления оптимальной обработки сигналов (создания оптимального фильтра) необходимо выполнить постановку, формализацию и решение задачи синтеза структуры фильтров, удовлетворяющих заданной совокупности показателей качества и ограничений. В настоящее время главным критерием при проектировании таких систем является минимизация среднеквадратичной ошибки. В зависимости от того, какими уравнениями описывается состояние системы, оптимальные (квазиоптимальные) фильтры подразделяются на линейные и нелинейные. Первые обычно базируются на оптимальном фильтре Калмана, а вторые – на многоканальных цифровых фильтрах, качество работы которых зависит от числа каналов. Современные вычислительные средства позволяют решать в режиме реального времени и задачи многомерной фильтрации, существенно более сложные, чем цифровая фильтрация одномерных сигналов, выполняемая с помощью сигнальных процессоров или многопроцессорных систем

Основное отличие вейвлет-фильтрации от традиционных методов выделения радиотехнических сигналов из шумов заключается в том, что выбор параметров вейвлет-фильтра довольно слабо зависит от характеристик спектра сигнала. Это позволяет избежать тех трудностей, которые обычно сопровождают выбор параметров АЧХ традиционного фильтра, когда слишком узкое частотное окно приводит к искажению формы сигнала и ухудшению разрешающей способности системы, а слишком широкое окно – к неэффективности процесса фильтрации из-за большого уровня шумов на выходе.

Следует отметить, что метод вейвлет-фильтрации сигналов обеспечивает глубокое подавление шума при сохранении исходной структуры сигнала. В целом же применение предложенного алгоритма выявило высокую эффективность вейвлет-фильтрации при обработке сигналов сложной формы, что объясняется основными свойствами самих базовых вейвлетов, а именно их ограниченностью автономностью. С его помощью, например, удалось успешно обработать искаженное шумом изображение модельной атомной решетки со случайным расположением дефектов (вакансий). Фурье-фильтрация в данном случае оказалась малоэффективной.

Наиболее часто при сжатии сигналов применяются фильтры с конечной импульсной характеристикой – один из видов линейных цифровых фильтров, характерной особенностью которого является ограниченность по времени его импульсной характеристики (с какого-то момента времени она становится точно равной нулю). Такой фильтр называют ещё нерекурсивным из-за отсутствия обратной связи. Знаменатель передаточной функции такого фильтра – константа. В таких фильтрах значение любого отсчета выходного сигнала определяется суммой масштабированных значений предыдущих отсчетов в силу свойства линейности. Можно сказать иначе: значение выхода фильтра в любой момент времени есть значение отклика на мгновенное значение входа и сумма всех постепенно затухающих откликов предыдущих отсчетов сигнала, которые всё ещё оказывают влияние на выход отсчетов. Процедура построения реального фильтра заключается в минимизации отклонения заданного числа свойств от свойств идеального фильтра в работе. При конструировании обычных и вейвлет-фильтров для этой минимизации используются различные критерии, следует отметить, что под обычными в данной работе понимаются фильтры, используемые в субполосном кодировании. Длина фильтра важна в силу двух причин. Во-первых, в таких приложениях, как сегментация изображения, длинный фильтр приведет к неверной локализации контуров, так как на протяжении одного фильтра могут встретиться два контура. Во-вторых, короткие фильтры вычислительно

экономнее. К симметричным четным фильтрам применимо полифазное построение, что снижает вычислительную сложность в два раза.

Использование методов вейвлет-обработки прочно вошло в основные направления обработки сигналов: удаление шума, сокращение избыточности представления сигналов и изображений (сжатие). Использование априорной информации о характере обрабатываемого сигнала позволяет синтезировать оптимизированные вейвлет-фильтры. Разработка новых методов представления вейвлет-фильтров позволяет гибче подойти к их синтезу.

При обработке реальных неоднородных и нестационарных сигналов возникают задачи адаптации рассмотренных процедур обнаружения и фильтрации многомерных сигналов. Для этих целей были разработаны достаточно простые безыдентификационные алгоритмы псевдоградиентной адаптации, описанные в работах. Областью применения таких алгоритмов является адаптивная поддержка весовых коэффициентов и пороговых уровней обнаружителей сигналов, декорреляция изображений, а также оценивание деформаций многомерных сеток.

Практическая реализация фильтрации сигнала, с применением адаптивной вейвлет-обработки в несколько этапов, при этом улучшение характеристик реального шума, а также увеличение отношения сигнал/шум при накоплении сигнала достигается применением разработанного медианного фильтра, параметры окна которого подбираются в зависимости от поведения эксцесса и симметрии.

Эффективность методики с использованием адаптивной вейвлет-фильтрации при обработке сигналов с помехой подтверждена практическим применением, в соответствии с работой, для обработки сигналов на частоте 4641 кГц в присутствии внешней помехи при добротности контура $Q = 200$. После предварительного анализа сигнала, полученного без образца сигнала, данные вводились в разработанную программу для последующего вейвлет-разложения и вычисления дисперсии и среднего, при этом введенные данные позволяют автоматизировать процесс выбора порога в каждой полосе разложения.

ЗАКЛЮЧЕНИЕ

Основные научные результаты диссертации

1 Разработан алгоритм работы кодера, выбраны средства реализации проекта, реализованы схемы декомпозиции и реконструкции сигнала в программной среде и произведено моделирование процесса кодирования-декодирования речевого сигнала. На выходе получился речевой кодер на основе алгоритма Малла с использованием трехуровневого преобразования. Проект реализован в среде Quartus и скомпилирован в микросхему XC6SLX100 семейства Spartan 6.

2 Проведено моделирование в среде Matlab с использованием реального речевого сигнала. В результате установлено практически полная идентичность преобразованного и восстановленного сигналов, что позволяет судить о высоком качестве восстановления. Точность восстановления составила величину не хуже 4 %.

3 В результате теоретических и экспериментальных исследований, а также математического и компьютерного моделирования описаны методы повышения эффективности вейвлет-преобразований при обработке, сжатии и восстановления радиотехнических сигналов.

4 Повышение эффективности методики очищения сигнала от шума и повышения точности диагностики полезного сигнала достигается применением двух подходов статистического анализа свойств экспериментального шумового сигнала. Основываясь на статистических свойствах исследуемого сигнала выявлены и рассмотрены два подхода с применением медианного фильтра.

Рекомендации по практическому использованию результатов

Разработанный модуль кодирования-декодирования может использоваться при создании высокоэффективных кодеков речевого сигнала для низкоскоростных линий коммуникаций.

Разработанные методы повышения эффективности применимы ко всем действующим кодерам, использующие в своей основе алгоритм Малла.

РЭЗІЮМЭ

Камінч Мікалай Юр'евіч

Павышэнне эфектыўнасці мадэлі маўленчага кодэра на аснове алгарытму Малах

Ключавыя словы: вэйвлетаў-Пераўтварэнне, алгарытм Малах, кампрэсія маўленчага сігналу, IP-тэлефанія.

Мэта работы: распрацоўка праграмнага забеспячэння для апрацоўкі і сціску галасавой інфармацыі, пабудаванага на базе ПЛІСА. Гэтак жа разгледжаны некалькі спосабаў паляпшэння эфектыўнасці мадэлі маўленчага кодэра на аснове алгарытму Малах.

Атрыманыя вынікі і іх навізна: абгрунтавана неабходнасць удасканалення метадаў і алгарытмаў нізкахуткасных кампрэсіі маўленчых сігналаў у рэальным часе і мэтазгоднасцю распрацоўкі прылад для нізкахуткасных кампрэсіі гаворкі. У сваю чаргу, для праектавання якасных нізкахуткасных вокодерных прылад неабходна выкарыстанне не толькі новых алгарытмаў сціску прамовы, але і адаптаваных для гэтых мэтаў апаратных сродкаў. Распрацавана некалькі алгарытмаў павышэння эфектыўнасці існуючых кодэраў маўленчых сігналаў.

Вобласць прымянення: IP-тэлефанія.

РЕЗЮМЕ

Коминч Николай Юрьевич

Повышение эффективности модели речевого кодера на основе алгоритма Малла

Ключевые слова: вейвлет-преобразование, алгоритм Малла, компрессия речевого сигнала, IP-телефония.

Цель работы: является разработка программного обеспечения для обработки и сжатия голосовой информации, построенного на базе ПЛИС. Так же рассмотрены несколько способов улучшения эффективности модели речевого кодера на основе алгоритма Малла.

Полученные результаты и их новизна: обоснована необходимость совершенствования методов и алгоритмов низкоскоростной компрессии речевых сигналов в реальном времени и целесообразностью разработки устройств для низкоскоростной компрессии речи. В свою очередь, для проектирования качественных низкоскоростных вокодерных устройств необходимо использование не только новых алгоритмов сжатия речи, но и адаптированных для этих целей аппаратных средств. Разработано несколько алгоритмов повышения эффективности существующих кодеров речевых сигналов.

Область применения: IP-телефония.

SUMMARY

Cominch Nikolai Yurievich

Improving the efficiency of the speech encoder model based on the Mall algorithm

Key words: wavelet transform, the Mallat algorithm, compression of the speech signal, IP-telephony.

The purpose of the work: is to develop software for processing and compression of voice information, built on the basis of FPGAs. Several ways to improve the efficiency of the speech encoder model based on the Mall algorithm are also considered.

The results and their novelty: the necessity of improving the methods and algorithms of low-speed compression of speech signals in real time and the feasibility of developing devices for low-speed speech compression. In turn, the design of high-quality low-speed vocoder devices requires the use of not only new algorithms for speech compression, but also adapted for this purpose hardware. Several algorithms for improving the efficiency of existing speech signal coders have been developed.

Scope: IP-telephony.