

Вашкевич М.И., Азаров И.С., Петровский А.А.

**КОСИНУСНО-МОДУЛИРОВАННЫЕ БАНКИ ФИЛЬТРОВ С
ФАЗОВЫМ ПРЕОБРАЗОВАНИЕМ: РЕАЛИЗАЦИЯ И
ПРИМЕНЕНИЕ В СЛУХОВЫХ АППАРАТАХ**

УДК 621.391+519.72

Р е ц е н з е н т ы:

заведующий кафедры информатики

Белорусского государственного университета информатики и радиоэлектроники

профессор, доктор технических наук А.А. Иванюк

заведующий кафедры интеллектуальных информационных технологий

Белорусского государственного университета информатики и радиоэлектроники

профессор, доктор технических наук В.В. Голенков

Вашкевич М.И., Азаров И.С., Петровский А.А., Косинусно-модулированные банки фильтров с фазовым преобразованием: реализация и применение в слуховых аппаратах // Вашкевич М.И., Азаров И.С., Петровский А.А. – Москва: Горячая линия «Телеком», 2014 – 219 с. : ил.

В монографии рассмотрены вопросы построения цифровых слуховых аппаратов. В общем виде описаны технические аспекты коррекции нарушений слуха. Большая часть материала посвящена разработке и эффективной реализации неравнополосных банков фильтров, играющих ключевую роль при проектировании цифровых слуховых аппаратов. Рассматриваются вопросы построения алгоритмов адаптивной компенсации акустической обратной связи, компрессии и шумоподавления, необходимых для реализации основных функций слухового аппарата. Монография представляет интерес для научных и инженерно-технических работников, специализирующихся в области проектирования устройств цифровой обработки сигналов, а также для аспирантов и студентов технических ВУЗов соответствующих специальностей.

*Рекомендовано советом учреждения образования «Белорусский государственный университет информатики и радиоэлектроники»,
протокол №4 от 20.12.2013 г.*

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ	8
ГЛАВА 1. ОБЩИЕ СВЕДЕНИЯ О СЛУХОВОЙ СИСТЕМЕ ЧЕЛОВЕКА И ОБЗОР ПОДХОДОВ К ПОСТРОЕНИЮ ЦИФРОВЫХ СЛУХОВЫХ АППАРАТОВ	11
1.1. Слух и понимание	11
1.2. Физиология уха человека	11
1.3. Психоакустические аспекты тугоухости	14
1.4. Функции слухового аппарата	16
1.4.1. Компрессия динамического диапазона	16
1.4.2. Шумоподавление	17
1.4.3. Подавление эффекта акустической обратной связи	18
1.5. Разработка слуховых аппаратов	21
1.5.1. Требования, предъявляемые к слуховым аппаратам	21
1.5.2. Современные подходы к построению слуховых аппаратов	23
1.5.3. Структура слухового аппарата с обработкой в перцептуальной частотной шкале	28
1.6. Обзор методов построения неравнополосных банков фильтров	30
1.7. Краткие выводы	34
ГЛАВА 2. СИНТЕЗ НЕРАВНОПОЛОСНОГО КОСИНУСНО-МОДУЛИРОВАННОГО БАНКА ФИЛЬТРОВ	35
2.1. Косинусно-модулированный банк фильтров (КМБФ)	35
2.1.1. Основные понятия	35
2.1.2. Полифазная реализация КМБФ	38
2.1.3. Эффективная реализация косинусной модуляции	40
2.2. Неравнополосный косинусно-модулированный банк фильтров	42
2.2.1. Фазовое преобразование	42
2.2.2. Деформация частотной оси КМБФ	44
2.2.3. Расчет коэффициента деформации частотной оси КМБФ для аппроксимации шкалы барков	46
2.2.4. Выбор коэффициентов децимации/интерполяции	48
2.2.5. Выравнивание групповой задержки	50

2.3. Метод расчета фильтра-прототипа неравнополосного косинусно-модулированного банка фильтров	53
2.3.1. Постановка задачи	53
2.3.2. Расчетные выражения	54
2.3.3. Расчет фильтра-прототипа НКМБФ	56
2.4. Результаты экспериментов	59
2.4.1. Расчет коэффициентов децимации/интерполяции	59
2.4.2. Оптимизация фильтра-прототипа	61
2.4.3. Бичастотная функция неравнополосного КМБФ	63

ГЛАВА 3. ПЕРЕКРЫВАЮЩИЕСЯ ПРЕОБРАЗОВАНИЯ И ИХ ПРИМЕНЕНИЕ В СЛУХОВЫХ АППАРАТАХ 67

3.1. Причины применения перекрывающихся преобразований в слуховых аппаратах	67
3.2. Перекрывающееся преобразование (ПП)	68
3.3. Дополненное модулированное ПП (ДМПП)	68
3.3.1. Модулированное перекрывающееся преобразование	69
3.3.2. Определение ДМПП	69
3.3.3. Схема компенсации акустической обратной связи на основе ДМПП	71
3.3.4. Предсказатель $\hat{H}[m]$	72
3.4. Модифицированное дискретное косинусное преобразование с деформацией частотной оси	74
3.4.1. Модифицированное дискретное косинусное преобразование (МДКП)	75
3.4.2. Искривление частотной оси МДКП	77
3.4.3. Применение МДКП с деформацией частотной оси в слуховом аппарате	77
3.5. Краткие выводы	80

ГЛАВА 4. СИСТЕМАТИЧЕСКИЙ МЕТОД СИНТЕЗА БЫСТРЫХ АЛГОРИТМОВ ВЫЧИСЛЕНИЯ ДИСКРЕТНОГО КОСИНУСНОГО ПРЕОБРАЗОВАНИЯ ПРОИЗВОЛЬНОГО РАЗМЕРА 81

4.1. Обоснование разработки систематического метода синтеза быстрых алгоритмов ДКП	81
4.2. Анализ методов синтеза быстрых алгоритмов дискретных преобразований	82
4.3. Синтез быстрых алгоритмов дискретных преобразований с использованием полиномиальной алгебры	84
4.3.1. Представление алгоритма	84
4.3.2. Полиномиальное преобразование	85
4.4. Алгебраический метод синтеза быстрых алгоритмов ДКП	87
4.4.1. Описание метода	87
4.4.2. Синтез быстрого алгоритма ДКП-4 формата 7	87
4.4.3. Рекурсивный быстрый алгоритм ДКП-4	93
4.4.4. Рекурсивный быстрый алгоритм ДКП-2	97
4.5. Основные результаты и выводы	101

ГЛАВА 5. АЛГОРИТМЫ ОБРАБОТКИ СИГНАЛА В СЛУХОВОМ АППАРАТЕ НА ОСНОВЕ НЕРАВНОПОЛОСНОГО БАНКА ФИЛЬТРОВ 102

5.1. Модификация спектра сигнала в слуховом аппарате	102
5.1.1. Принцип работы	103
5.1.2. Шумоподавление	106
5.1.3. Компенсация потери слуха	111
5.1.4. Компрессия динамического диапазона	113
5.2. Компенсация эффекта акустической обратной связи с использованием неравнополосного банка фильтров	118
5.2.1. Введение	118
5.2.2. Метод субполосной компенсации эффекта акустической обратной связи	120
5.2.3. Оценка эффективности компенсации сигнала АОС	125
5.3. Основные результаты и выводы	131

ГЛАВА 6. РЕАЛИЗАЦИЯ СЛУХОВОГО АППАРАТА НА МОБИЛЬНОЙ ВЫЧИСЛИТЕЛЬНОЙ ПЛАТФОРМЕ 132

6.1. Преимущества и недостатки мобильных платформ для реализации слуховых аппаратов	132
6.2. Схема обработки сигнала для коррекции слуха	135
6.3. Определение целевых параметров звукоусиления	140
6.3.1. Аудиограмма	140
6.3.2. Расчет целевого усиления и коэффициентов корректирующего фильтра	142
6.3.3. Компенсация нарушения ощущения громкости	145
6.3.4. Временные параметры компрессора динамического диапазона	147
6.4. Реализация слухового аппарата на мобильной платформе	150
6.4.1. Схема включения приложения	150
6.4.2. Полоса пропускания	151
6.4.3. Алгоритм обработки сигнала в слуховом аппарате на мобильной платформе	151
6.4.4. Задержка обработки сигнала	157
6.5. Результаты экспериментов	157
6.5.1. Конфигурация устройства	157
6.5.2. Полное акустическое усиление	158
6.5.3. Пример обработки сигнала	158
6.6. Резюме	160

ГЛАВА 7. АППАРАТНАЯ РЕАЛИЗАЦИЯ ВЫЧИСЛИТЕЛЬНЫХ ЭЛЕМЕНТОВ СЛУХОВОГО АППАРАТА НА ОСНОВЕ НЕРАВНОПОЛОСНОГО БАНКА ФИЛЬТРОВ	161
7.1. Структура процессора слухового аппарата	161
7.2. Реализация неравнополосного КМБФ	163
7.3. Цепочка фазовых звеньев	164
7.4. Блок полифазных компонент фильтра-прототипа	168
7.5. Блок косинусной модуляции	170
7.5.1. Быстрый алгоритм ДКП с использованием техники кодирования алгебраическими числами	171
7.5.2. Реализация 16-точечного ДКП-2	171
7.5.3. IP-ядра быстрого алгоритма ДКП	177

7.6. Система быстрого прототипирования неравнополосного косинусно-модулированного банка фильтров	182
7.7. Основные результаты и выводы	186
ЗАКЛЮЧЕНИЕ	187
ЛИТЕРАТУРА	189
ПРИЛОЖЕНИЕ А. Краткие сведения из теории групп и теории Галуа	205
ПРИЛОЖЕНИЕ Б. Полиномы Чебышева	207
ПРИЛОЖЕНИЕ В. Кохлеарный банк фильтров	208

ВВЕДЕНИЕ

По данным всемирной организации здравоохранения на 2005 г. около 278 млн человек в мире имели умеренные или серьезные нарушения слуха [1]. 80% из них живет в странах с низким и средним уровнем дохода.

Распространенными причинами нарушения слуха являются инфекционные заболевания (например менингит, корь и др.), воздействие чрезмерного шума, травмы, старение и применение ототоксичных лекарственных средств [2]. Большой частью слуховой дисфункцией страдают люди в трудоспособном возрасте. Поэтому возвращение или улучшение слуха таким людям – так называемая слуховая реабилитация – представляет собой важную медико-социально-экономическую проблему. Важное место в общем комплексе мер по слуховой реабилитации занимает применение специальных технических средств – слуховых приборов, однако с помощью производимых в настоящее время слуховых аппаратов можно удовлетворить менее 10% глобальных потребностей. В развивающихся странах слуховые аппараты имеют менее 1 из каждых 40 человек нуждающихся в них. В то же время возможности слуховых аппаратов постоянно расширяются благодаря совместным усилиям ученых из различных областей науки, включая физиологию, аудиологию, психо- и биофизику и отоларингологию.

Все полученные знания о различных патологиях и возможных способах коррекции слуха используются слуховыми аппаратами в виде соответствующих алгоритмов цифровой обработки сигналов (ЦОС). Однако из-за особенностей приложения разработка и реализация таких алгоритмов сопряжена с серьезными сложностями. Во-первых, по своему назначению слуховой аппарат должен выполнять непрерывный ввод, вывод и обработку звукового сигнала; во-вторых, задержка сигнала между входом и выходом аппарата (алгоритмическая задержка обработки в прямом канале аппарата) не должна превышать 7–9 мс [3–6]; в-третьих должно быть обеспечено крайне низкое энергопотребление устройства, поскольку слуховой аппарат выполняется в виде миниатюрного изделия и не может иметь мощного автономного источника энергии. Перечисленные особенности предъявляют специальные требования как к алгоритмам обработки, так и к вычислительной платформе устройства.

В отношении алгоритмов обработки наиболее важными их характеристиками в данном контексте являются эффективность реализации (вычислительная сложность) и алгоритмическая задержка. Центральным компонентом цифровой обработки звукового сигнала является частотно-временное преобразование. Благодаря специфике приложения (непрерывный ввод-вывод) наиболее подходящей является реализация преобразования в виде банка фильтров [7]. Причем предпочтительно, чтобы сигнал представлялся в частотной области с разным разрешением, учитывающим особенности слухового восприятия человека [3]. Само по себе проектирование таких банков фильтров с заданными параметрами является актуальной задачей для многих приложений, применительно же к слуховому аппарату эта задача приобретает особую важность и потому в настоящей работе ей уделяется особое внимание. Кроме того, на основе частотной декомпозиции сигнала, выполняемой банком фильтров, предлагаются специальные, адаптированные к требованиям приложения, алгоритмы подавления шума, обратной акустической связи, компрессии динамического диапазона и др.

Что касается вычислительной платформы, то основными требованиями являются производительность, низкое энергопотребление, низкая тактовая частота и архитектура, позволяющая эффективно выполнять требуемые алгоритмы обработки. Сегодня существует две основные альтернативы: процессоры цифровой обработки сигналов (ПЦОС) [8,9] и программируемые логические интегральные схемы (ПЛИС). Преимуществом ПЦОС является их высокая производительность, наличие аппаратных ускорителей многих операций, а также скорость разработки и отладки программного кода. Однако главным недостатком ПЦОС для их использования в качестве вычислителей для слуховых аппаратов является тот факт, что им характерна высокая тактовая частота. Это в свою очередь приводит к повышенному энергопотреблению, что не соответствует требованиям целевого приложения. Необходимо учитывать, что алгоритмы обработки имеют высокую степень параллелизма, что делает возможным их эффективную реализацию на ПЛИС в виде параллельно работающих структур. Использование параллельных структур позволяет уменьшить тактовую частоту процессора слухового аппарата, поскольку данные в этом случае будут обрабатываться в темпе их поступления.

При проектировании слухового аппарата желательно использовать высокоуровневые структуры и блоки обработки, обеспечивающие гибкость и возможность быстрой адаптации платформы к новым требованиям и алгоритмам. Новые требования возникают вследствие развития технических средств, с возникновением более производительных устройств с низким потреблением или при появлении и диагностике новых патологий слуха. Поэтому в настоящей работе предлагается процессор слухового аппарата в виде высокоуровневой реконфигурируемой структуры, объединяющей функциональные макроблоки. В процессе разработки слухового аппарата данные макроблоки настраиваются и объединяются в соответствии с заданными требованиями, после чего процессор синтезируется на ПЛИС. Этим достигается сокращение производственного цикла изделия (быстрое прототипирование) с сохранением его гибкости и функциональности.

В монографии приводится пример построения слухового аппарата на базе современной мобильной вычислительной платформы (iPhone). Интерес к такого рода реализации обусловлен широким распространением высокопроизводительных смартфонов, позволяющих выполнять обработку сигнала в режиме реального времени. При разработке слухового аппарата на базе мультимедийной вычислительной платформе необходимо учитывать особенности ее архитектуры для того, чтобы обеспечить требуемую функциональность. В работе предлагается специальная, адаптированная для смартфона, схема обработки сигнала с малой групповой задержкой и пониженной вычислительной сложностью.

ГЛАВА 1

ОБЩИЕ СВЕДЕНИЯ О СЛУХОВОЙ СИСТЕМЕ ЧЕЛОВЕКА И ОБЗОР ПОДХОДОВ К ПОСТРОЕНИЮ ЦИФРОВЫХ СЛУХОВЫХ АППАРАТОВ

1.1. Слух и понимание

Слух и понимание представляют собой два разных аспекта восприятия акустических сигналов. Для того чтобы акустически сигнал (речь) стал понятным, необходимо, чтобы он был слышимым [10]. Если сигнал не слышен, его невозможно понять, однако обратное неверно. Тот факт, что сигнал является слышимым, вовсе не означает, что он будет понят. Таким образом, «понятность» или «разборчивость» зависят от того, может ли мозг человека распознать слышимые звуковые сигналы. Часто случаются ситуации, в которых человек может слышать акустический сигнал (например, разговор двух людей, находящихся по отдалу), однако не понимать информации, которую несет этот сигнал. Таким образом, между слышимостью и разборчивостью есть существенное различие. Тем не менее, эти термины часто используются как синонимы.

С проблемой «непонимания» сталкиваются слабослышащие люди (в особенности те, у которых имеются патологии внутреннего уха). Они оказываются в такой ситуации, когда не способны расшифровать звуковые сигналы с помощью доступной акустической информации.

1.2. Физиология уха человека

Ухо состоит из трех основных частей (рис. 1.1):

- наружное ухо, состоящее из расположенной вне черепа ушной раковины и слухового прохода, который заканчивается барабанной перепонкой;
- заполненное воздухом среднее ухо, состоящее из барабанной пере-

- понки и слуховых косточек – молоточка, наковальни и стремечка¹;
- внутреннее ухо, или улитка, заполненное жидкостью. Оно преобразует механические колебания в электрические сигналы.



Рис. 1.1. Анатомия уха человека

Звуковой сигнал достигает барабанной перепонки, расположенной в конце слухового прохода. Изменения давления у барабанной перепонки очень малы, но она столь чувствительна, что даже небольшие колебания давления у поверхности перепонки приводят к достаточному её смещению. От барабанной перепонки через слуховые косточки звуковые волны преобразуются в механические колебания и передаются к овальному окну улитки. Давление, достигающее улитки, в 20–25 раз превышает давление, приложенное к барабанной перепонке. В улитке давление преобразуется в электрические сигналы, воспринимаемые мозгом [10].

Улитка – самая сложная часть уха. Физически она представляет собой заполненную жидкостью полость, по форме напоминающую раковину улитки. Высота улитки около 5 мм, количество витков – чуть больше трех. Если бы улитку можно было развернуть, её длина составила бы около 35 мм. Внутри улитки находится мембрана, называемая основной или базилярной (рис. 1.2). Она прикрепляется к боковым сторонам улитки, оставаясь при этом незначительно натянутой. Проникающая через овальное

¹Стремечко через овальное окно соединяется с внутренним ухом

окно в улитку звуковая энергия изгибает основную мембрану, генерируя распространяющуюся по ней бегущую волну. У овального окна мембрана достаточно узкая и жесткая. По мере приближения к вершущке она становится более широкой и гибкой. Движения овального окна приводят к образованию волны, амплитуда которой постепенно нарастает вплоть до резонансного пика в некоторой точке основной мембраны, после чего волна быстро затухает. В результате такого строения каждый участок мембраны имеет свою резонансную частоту. Так, высокочастотный звук затрагивает волосковые клетки, располагающиеся в начале улитки, а звуковые колебания низкой частоты – окончания в ее вершине. В целом, можно представить улитку как набор пространственных фильтров. Диапазон воспринимаемых ухом частот составляет от 20 Гц до 20 кГц.

На основной мембране находится структура, называемая кортиевым органом. Он отвечает за преобразование механической энергии резонансной волны в электрические сигналы, распознаваемые мозгом. Основным компонентом кортиева органа являются волосковые клетки, играющие главную роль в преобразовании энергии волны в электрические стимулы. Улитка содержит около 4000 внутренних волосковых клеток (ВВК) и порядка 12000 наружных волосковых клеток (НВК). Они организованы в группы (по три ВВК на каждую НВК), расположенные на расстоянии 10 мкм друг от друга. Необходимо отметить функционально различие двух типов волосковых клеток, играющих важную роль в процессе звуковосприятия [11]. ВВК представляют собой простые рецепторы, которые передают в мозг информацию в виде электрических импульсов. В свою очередь НВК играют

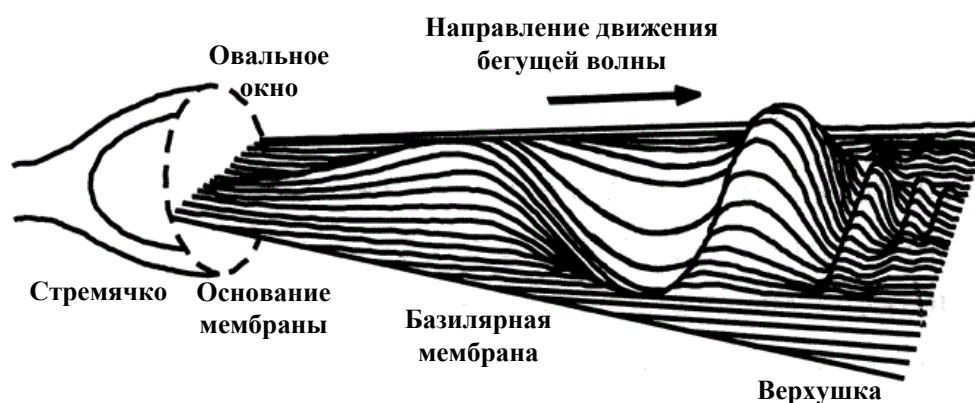


Рис. 1.2. Основная мембрана

«активную» роль в анализе звуков во внутреннем ухе. НВК под управлением импульсов получаемых от головного мозга изменяют свою длину. Это влияет на колебания основной мембраны, усиливая или подавляя их. Благодаря такой способности нелинейной «подстройки» человек может слышать как очень громкие, так и очень тихие звуки с высокой степенью дифференциации уровней и частот акустических сигналов. Таким образом, внутреннее ухо – не просто пассивный приемник, а активный усилитель.

1.3. Психоакустические аспекты тугоухости

Потерей слуха в различных формах страдают многие люди. Существует несколько причин ухудшения слуха [2, 12]. Во-первых, с возрастом часто в ухе происходят патологические изменения, приводящие к понижению остроты слуха (главным образом к высоким частотам). Во-вторых, к ослаблению слуха приводят различные заболевания уха, такие как катар, отосклероз или осложнения на ухо, вызванные инфекционными болезнями. Снижает слух также длительное воздействие шума и регулярное прослушивание музыки при высокой интенсивности звука (более 100 дБ). В результате действия указанных причин у человека снижается разборчивость речи.

Рассмотрим проблему нарушения слуха с точки зрения того, как слуховой анализатор обрабатывает звук и как влияет поражение внутреннего уха на восприятие. Как уже говорилось, благодаря особой механике внутреннее ухо способно осуществлять очень точный анализ частотного состава звуковых сигналов. Иными словами, оно может точно определить соотношение высоких и низких частот в акустическом сигнале. При его поражении страдают как пассивная, так и активная функции волосковых клеток.

При улитковой тугоухости серьезно страдают два процесса – анализ частоты звука и способность различать уровни сигнала, что существенно снижает возможность слышать и понимать услышанное. Причина заключается в поражении наружных волосковых клеток (НВК). В результате утрачивается очень важная для нормального слуха функция – активное усиление во внутреннем ухе. Часто для тугоухих людей элементы речи,

отличаемые по частоте ([о] – низкочастотный, [и] – высокочастотный) становятся неразличимыми. Рассмотрим, например, слова «кот» и «кит». Единственное их различие – частотный состав гласного звука. Если частотный анализатор человека перестал работать нормально, различить эти слова будет непросто.

Второй важный аспект заключается в восприятии физических уровней сигнала. Тугоухий человек не слышит тихих звуков, но громкие звуки представляются им такими же, как и людям с нормальным слухом. Это явление носит название «рекруитмент». Человек с нормальным слухом спокойно воспринимает звуки интенсивностью от 10 дБ (очень тихо) до 100 дБ (очень громко). При тугоухости этот диапазон существенно сужается и может составлять, например, от 50 дБ (очень тихо) до 100 дБ (очень громко). В результате восприятие громкости меняется гораздо быстрее, чем в норме. Небольшое изменение уровня звука приводят к существенному изменению уровня воспринимаемой громкости. Человек с нормальной слуховой системой может представить себе этот эффект как появление внезапного громкого шума.

Ещё одно последствие тугоухости, которое оказывает влияние на разборчивость, состоит в неадекватном восприятии изменений громкости акустических сигналов, например речи или музыки. Между тем, колебания уровня звука содержат очень важную информацию. Тихие компоненты речевого сигнала – тихие гласные или согласные – становятся неслышными, в результате чего затрудняется распознавание промежутков между словами. С другой стороны, звуки высокой интенсивности воспринимаются как слишком громкие, что в итоге приводит к значительным искажениям акустических сигналов. Диапазон громкости обычных звуков, описываемый слышащим человеком как «умеренный» или «приятный», превращается в узкий промежуток между «очень тихо» и «очень громко».

1.4. Функции слухового аппарата

1.4.1. Компрессия динамического диапазона

Основной задачей слухового аппарата является компенсация нарушения слуховой функции [2]. Как объяснялось выше, тугоухий человек, как правило, имеет повышенный порог слышимости, который мешает воспринимать звуки малой интенсивности. Тем не менее, восприятие звуков высокой интенсивности близко к тому, как их воспринимают люди с нормальным слухом. Таким образом, одной из главных функций слухового аппарата является отображения широкого динамического диапазона речевого сигнала в суженный динамический диапазон остаточного слуха [3]. Более того, нарушение слуха, как правило, имеет частотно-зависимый характер [13], поэтому различная степень компенсации требуется для различных частотных диапазонов (см. рис. 1.3). Решением данной проблемы является построение многоканальной системы, такой как банк фильтров, с различной степенью компенсации в каждом канале [14]. Устройство, выполняющее описанную обработку сигнала, называется многоканальным компрессором динамического диапазона.

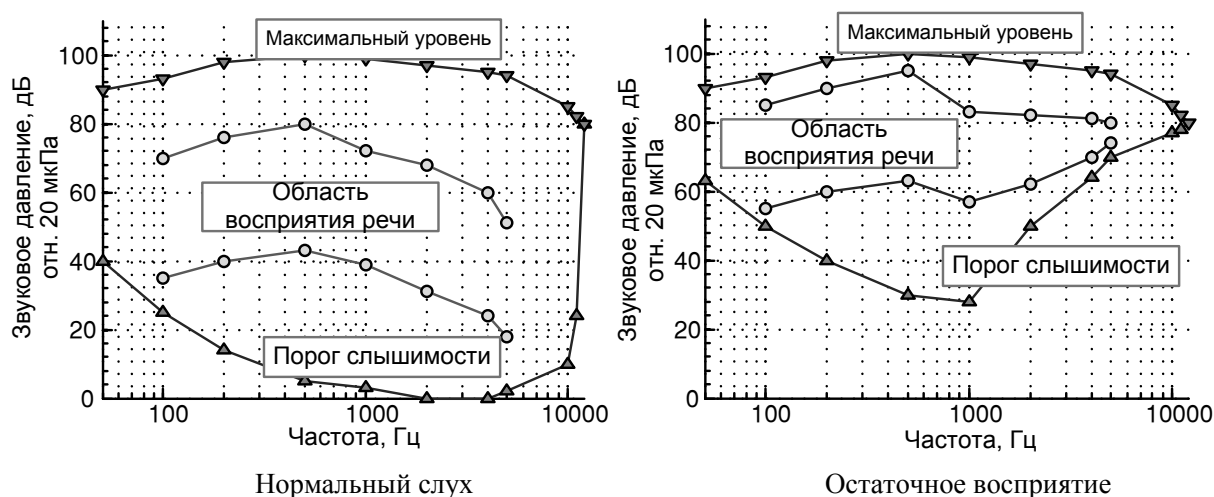


Рис. 1.3. Динамический диапазон звуков, воспринимаемый нормальным и тугоухим человеком

Одной из проблем, связанной с проектированием многоканального компрессора, является согласование частотного разрешения цифровой системы со слуховой системой человека. Так, например, многие процедуры

подстройки слухового аппарата, основанные на корректировке громкости в поврежденном ухе [15], предполагают выполнение частотного анализа, согласованного со слуховой системой. Частотный анализ, используемый в цифровой обработке сигналов, такой как дискретное преобразование Фурье, имеет равномерное разрешение во всем частотном диапазоне. Частотное же разрешение слуховой системы человека более точно моделируется неравнополосным банком фильтров. Исходя из того, что слуховая система человека более чувствительна к изменениям сигнала в области нижних частот, банк фильтров должен обладать следующим свойством: ширина полос должна возрастать с увеличением частоты [16].

1.4.2. Шумоподавление

Другой важной функцией слухового аппарата является шумоподавление. Фоновый шум значительно ухудшает разборчивость речи, особенно для людей с тугоухостью [17]. Задача системы шумоподавления ослабить фоновый шум и повысить качество речевого сигнала в сложной акустической обстановке. Повышение разборчивости речи выполняется за счет увеличения соотношения сигнал/шум без внесения искажения в полезный сигнал. В зависимости от числа микрофонов системы шумоподавления делятся на одноканальные и многоканальные [18]. Одноканальные алгоритмы шумоподавления широко изучены. Грубо эти алгоритмы можно классифицировать на параметрические и непараметрические. Алгоритмы одноканального шумоподавления разрабатываются для очистки зашумленных речевых сигналов, используя только один микрофон, т.е. основываясь только на временных и спектральных различиях между речевым и шумовым сигналами. Одноканальное шумоподавление становится особенно сложной задачей в случае нестационарного источника шума при низком соотношении сигнал/шум. Поскольку речь и шум, как правило, занимают одни и те же частотные полосы, одноканальные системы шумоподавления обычно вносят слышимые артефакты и искажения в полезный речевой сигнал.

Многие недостатки одноканальных систем шумоподавления можно устранить, если установить дополнительные микрофоны [19]. Благодаря миниатюризации микрофонов слуховой аппарат в настоящее время может

быть оснащен двумя или даже тремя микрофонами. Обычно, источник полезного сигнала и источник шума имеют различное расположение. Это пространственное разделение можно использовать для получения лучших спектральных оценок фонового шума и полезного речевого сигнала [20]. Многоканальные системы шумоподавления часто основываются на методе обобщенного подавителя бокового лепестка [21] (GSC – generalized sidelobe canceller) или технике многоканальной винеровской фильтрации [22].

В настоящее время все большее распространение получает так называемое открытое слухопротезирование (open fitting), при котором в слуховой проход пациента не помещают вкладыша. Это приводит к возрастанию уровня комфорта при использовании слухового аппарата и позволяет избежать эффекта окклюзии². Главный недостаток открытого слухопротезирования – усиление эффекта от проникновения необработанного акустического сигнала в ушной канал пациента. Обычные схемы шумоподавления не могут учесть и компенсировать данный эффект. Более того, в такой ситуации применение схемы шумоподавления перестает давать эффект, поскольку необработанный («зашумленный») сигнал свободно поступает на барабанную перепонку. Для препятствия проникновению нежелательного шумового сигнала необходимо использовать активную схему компенсации шума (ANC – active noise control) [23]. В данном подходе схема генерирует акустическую «антиволну», которая создает эффект компенсации сигнала в определенной области пространства (т.н. «зона тишины»). В слуховых аппаратах часто в один модуль объединяются шумоподавление и активная схема компенсации шума (АСКШ) [24]. При этом стремятся при помощи АСКШ создать зону тишины (zone-of-quiet) в ушном канале в области барабанной перепонки.

1.4.3. Подавление эффекта акустической обратной связи

Стремление сделать слуховой аппарат менее заметным приводит к тому, что размеры современных слуховых аппаратов постоянно уменьшаются. К сожалению, это обстоятельство ведет к появлению новых нежелательных эффектов, таких как акустическая обратная связь. Акустическая обратная

²См. описание эффекта окклюзии на стр 19.

связь (АОС) обусловлена близким расположением динамика и микрофона (рис. 1.4).



Рис. 1.4. Акустическая обратная связь

Эффект акустической обратной связи проявляется, когда усиленный слуховым аппаратом звук попадает в микрофон, создавая замкнутый контур. АОС возникает из-за отверстия в ушном вкладыше, который «закрывает» слуховой аппарат. Отверстие в ушном вкладыше необходимо для уменьшения эффекта окклюзии [15, 25]. В аудиологии этот термин относится к явлению, которое заключается в том, что при закрытии отверстия ушного прохода громкость низкочастотных звуков увеличивается благодаря специфической звуковой проводимости костей. Часто эффект окклюзии возникает, когда человек начинает носить слуховой аппарат, поскольку ушной вкладыш (или корпус внутриканального аппарата) плотно закрывает ушной проход. Вследствие эффекта окклюзии человек слышит свой голос непривычно глухим и искаженным. Кроме того, становятся сильнее слышны звуки тела, например, при жевании или ходьбе. Чтобы уменьшить эффект окклюзии, во вкладыше или в корпусе внутриканального аппарата проделывается отверстие (vent). Полностью устраняет эффект окклюзии открытый ушной вкладыш, однако это значительно усугубляет проблему АОС и ограничивает максимальный допустимый коэффициент усиления слухового аппарата. Внешне эффект АОС проявляется как продолжительное, высокочастотное колебание, воспроизводимое слуховым аппаратом.

Существуют два подхода к устранению эффекта АОС: подавление прямой связи (feedforward suppression) и компенсация обратной связи (feedback cancellation). Оба подхода преследуют общую цель – увеличить коэффициент усиления слухового аппарата, минимизируя при этом артефакты, вносимые алгоритмом, такие как искажения, звон, «завывание». Проблема АОС может быть решена либо посредством полного устранения акустической связи между динамиком и микрофоном, либо путем удаления высокочастотных осцилляций из сигнала динамика. Наиболее широко изучены методы подавления «завывания» (howling suppression) на основе режекторных фильтров и адаптивные методы компенсации обратной связи.

Подавление прямой связи. В методах подавления прямой связи эффект АОС стараются ослабить в прямом канале слухового аппарата. Для этого часто используют режекторные фильтры с переменным коэффициентом усиления в узких частотных полосах, на которых проявляется эффект АОС [26, 27]. Данной технике присущи некоторые недостатки. Так, уменьшение коэффициента усиления на некоторых частотах потенциально может привести к ухудшению качества речи или даже уменьшить её разборчивость, если частоты режекции фильтров выбраны неправильно. Более того, описанный метод имеет реакционную природу, это означает, что режекторные фильтры начинают работать только после того, как АОС обнаружена детектирующим алгоритмом. Таким образом, техника подавления прямой связи не может обеспечить полного устранения эффекта АОС.

Компенсация обратной связи. В последнее время большее внимание уделяется алгоритмам компенсации обратной связи [28–30]. Целью данных алгоритмов является получение модели пути АОС, которая используется для предсказания сигнала обратной связи. Предсказанный сигнал затем вычитается из сигнала микрофона. Получающийся в результате сигнал содержит только полезную составляющую (при условии, что параметры модели пути АОС оценены достаточно точно). В идеальном случае описанный подход сохраняет полезный сигнал на входе прямого канала слухового аппарата. Поскольку путь АОС может быстро изменяться в зависимости от акустической среды, то он, как правило, моделируется при помощи адап-

тивного КИХ-фильтра. Путь АОС включает медленно изменяющуюся часть (микрофон, усилитель, динамик) и быстро изменяющуюся часть (акустика отверстия в ушном вкладыше, внешняя акустическая среда). Указанные факторы в значительной степени влияют на характеристики пути АОС [31]. Примером влияния внешних факторов может служить поднесенный к уху телефон [32], шляпа, одеваемая пользователем слухового аппарата, или реверберация комнаты [33]. Таким образом, точность оценки пути АОС является ключевым моментом, поскольку несоответствие оцененного пути реальному ведет к нестабильности слухового аппарата и появлению искажений в выходном сигнале.

Есть две разновидности алгоритмов компенсации АОС: с прерывной и непрерывной адаптацией [26]. В первом случае коэффициенты фильтра начинают адаптироваться только после того, как обнаружена нестабильность. Эта особенность определяет реакционный характер этих алгоритмов – слуховой аппарат на некоторое время перестает быть стабильным перед тем, как активируется алгоритм компенсации, что в общем случае является нежелательным. Алгоритмы компенсации АОС с непрерывной адаптацией имеют упреждающий характер – критические частоты, на которых проявляется эффект АОС, выявляются до момента возникновения «завываний». Оценку акустического пути можно выполнить, используя стандартные алгоритмы адаптивной фильтрации на основе метода наименьших средних квадратов. Однако использование непрерывной адаптации наталкивается на проблему корреляции между сигналом динамика и сигналом микрофона, которая возникает из-за присутствия замкнутого контура (сигнал передается от динамика на микрофон). Эта проблема трудно устранима в случае спектрально окрашенных сигналов, таких как музыка и речь.

1.5. Разработка слуховых аппаратов

1.5.1. Требования, предъявляемые к слуховым аппаратам

Несмотря на непрерывное развитие полупроводниковых технологий и увеличение вычислительной мощности современных процессоров цифровой обработки сигналов проектирование слухового аппарата остается весьма

сложной задачей [34]. В первую очередь это связано с техническими требованиями, предъявляемыми к слуховым аппаратам. К наиболее важным требованиям можно отнести следующие:

- обработка сигнала в реальном времени;
- низкое энергопотребление;
- малая алгоритмическая задержка.

Рассмотрим указанные требования более подробно. Необходимость обработки сигнала в реальном масштабе времени очевидна. Человек с ослабленным слухом желает воспринимать звуковые явления, относящиеся к настоящему моменту времени. Данное требование накладывает ограничения на методы и алгоритмы обработки сигнала в слуховом аппарате.

Требование низкого энергопотребления можно проиллюстрировать следующими доводами [7]. Слуховой аппарат должен обеспечивать 16-часовую бесперебойную работу в течение 7 дней (т.е. 112 часов). Если при этом использовать батарею на 1,4 В (180 мА×ч), то максимальный ток должен составлять 1,60 мА. Таким образом, нетрудно определить, что потребляемая слуховым аппаратом мощность составляет 2,24 мВт. Такое ограничение по потребляемой мощности ведет к ужесточению требований к вычислительной платформе. В частности, во всех практических реализациях для уменьшения потребляемой мощности тактовая частота работы процессора слухового аппарата колеблется в пределах 1–2,5 МГц [7, 35]. Кроме уменьшения частоты работы процессора проблему энергопотребления можно решать путем уменьшения вычислительной сложности алгоритмов, применяемых в слуховых аппаратах. Один из возможных путей состоит в уменьшении частоты дискретизации входного сигнала. В этом случае существенно сокращается число отсчетов, которые необходимо обработать в единицу времени. В настоящее время частота дискретизации, используемая в слуховых аппаратах, колеблется в пределах от 11 до 20 кГц [36]. Однако понижение частоты дискретизации ведет к тому, что многие полезные звуковые сигналы, которые имеют широкий частотный диапазон, перестают воспроизводиться слуховым аппаратом. Поэтому более разумный путь уменьшения энергопотребления состоит в разработке быстрых алгоритмов, обладающих минимальной вычислительной сложностью [37].

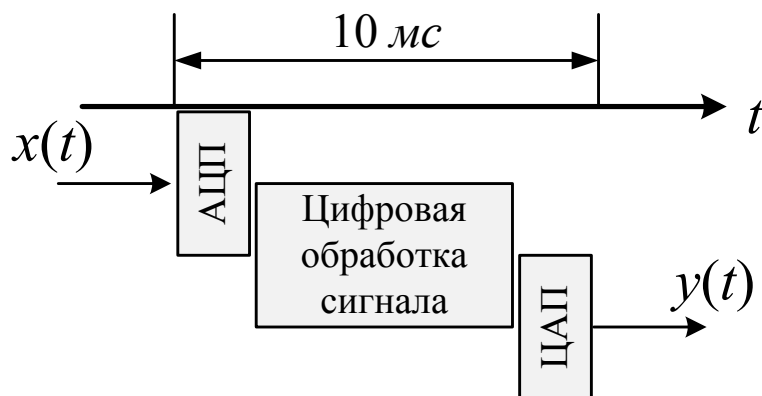


Рис. 1.5. Формирование задержки сигнала в слуховом аппарате при работе в реальном масштабе времени

Требование малой алгоритмической задержки слухового аппарата имеет несколько аспектов. В [5] указывается, что для предотвращения эффекта гребенки фильтров (*comb filter effect*) общее время на обработку сигнала в цифровом слуховом аппарате должно составлять менее 10 мс (рис. 1.5). Эффект гребенки фильтров вызывается суперпозицией обработанного (задержанного) сигнала с необработанным сигналом, который попадает в ушной канал напрямую (через отверстие в слуховом аппарате). Это составляет особенно серьезную проблему для слуховых аппаратов заушного типа. В общее время обработки сигнала в СА включает время на преобразование сигнала из аналоговой формы в цифровую, а также время на преобразование цифрового сигнала в аналоговый. С точки зрения проектирования слухового аппарата это означает, что время на цифровую обработку сигнала не должно превышать 6–8 мс [38]. В случае, если время обработки сигнала будет превышать 10 мс человек будет слышать эхо, создаваемое сигналом от слухового аппарата, который накладывается на барабанной перепонке с сигналом, поступающим напрямую из внешней среды. Рассмотренные требования являются ограничивающими факторами, и их необходимо учитывать на всех этапах разработки схемы слухового аппарата.

1.5.2. Современные подходы к построению слуховых аппаратов

В разделе рассматриваются подходы к построению слуховых аппаратов. В качестве классификационного признака выбран метод частотного

анализа, используемый в прямом канале слухового аппарата.

Слуховой аппарат с обработкой в частотной области В [39–41] предлагаются схемы обработки сигнала в слуховом аппарате, основанные на дискретном преобразовании Фурье (ДПФ). Общий принцип работы слуховых аппаратов на основе ДПФ можно пояснить, используя схему на рис. 1.6.

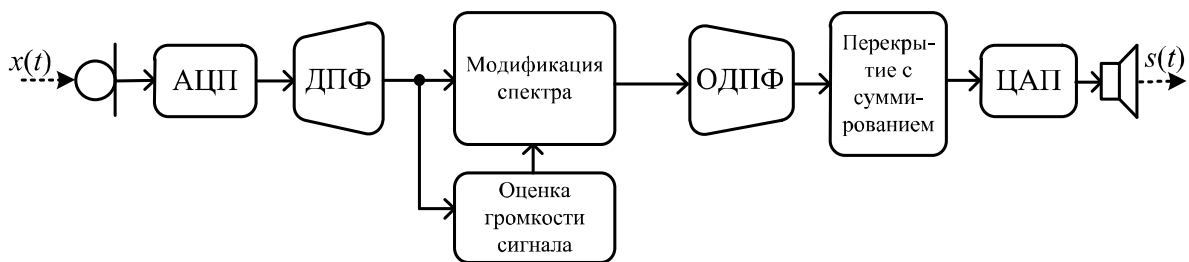


Рис. 1.6. Структура слухового аппарата на основе ДПФ

Входной сигнал $x(t)$ дискретизируется и разбивается на сегменты, которые переводятся в частотную область при помощи ДПФ. Амплитудный спектр сигнала изменяется в зависимости от текущей громкости звука и аудиограммы пользователя слухового аппарата. Модифицированный спектр сигнала затем преобразуется обратно во временную область посредством обратного ДПФ (ОДПФ) и реконструируется с использованием метода перекрытия с суммированием. Таким образом, входной сигнал усиливается нелинейно в зависимости от частоты и интенсивности звука.

Преимуществом подхода с использованием ДПФ заключается в хорошо разработанной теории синтеза быстрых алгоритмов вычисления ДПФ – БПФ [37, 42]. Также большое число современных процессоров цифровой обработки сигналов содержит в своем составе вычислительное ядро, реализующее БПФ. Однако подход с использованием ДПФ имеет алгоритмическую задержку порядка 14–25 мс, что как указывалось ранее является слишком большим значением. Кроме того частотный анализ на основе ДПФ не согласован со слуховой системой человека. Эти недостатки обусловили поиск и разработку альтернативных методов обработки сигнала в слуховых аппаратах.

Применение банков фильтров для обработки сигнала в слуховом аппарате. В работах [7,43] предлагается подход к построению слухового аппарата с использованием системы анализа/синтез на основе банка фильтров WOLA (*Weighted OverLap-Add*). Общая схема слухового аппарата на основе банка фильтров WOLA приведена на рис. 1.7 [44]. Структура банка фильтров WOLA состоящая из трех блоков: анализ, усиление и синтез показана на рис. 1.8.

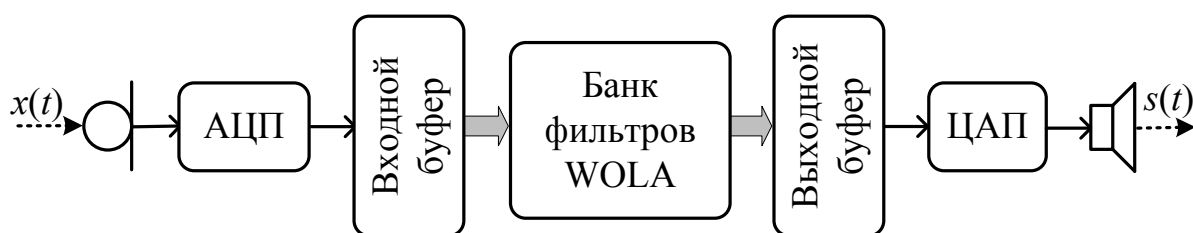


Рис. 1.7. Структура слухового аппарата на основе банка фильтров WOLA

На этапе анализа входная последовательность умножается на окно Хемминга длины L , что позволяет достигнуть хорошего частотного разрешения. На каждом шаге обрабатывается R текущих отсчетов (текущий блок) и $L - R$ предыдущих отсчетов (L/R блоков). Банк фильтров WOLA использует N частотных отсчетов БПФ взвешенного на окно входного сигнала. Поскольку выбрано четное выравнивание каналов банка фильтров, то в результате получается $N/2$ частотных полос. По этой причине на рис. 1.8 БПФ вычисляется для $N/2$ точек. Коэффициент передискретизации равен $O = N/R$. Передискретизация необходима, чтобы выполнить усиление частотных полос в банке фильтров без риска возникновения компонент наложения спектра. В блоке усиления происходит умножение $N/2$ частотных полос для коррекции слуха пользователя слухового аппарата. В блоке синтеза выполняется реконструкция модифицированного сигнала.

Типичные значения параметров банка фильтров WOLA приведены ниже [44]:

- частота дискретизации 16 кГц;
- размер входного блока $R = 64$ отсчета;
- длина окна анализа $L = 256$ отсчетов;
- формат БПФ $N/2 = 64$ точек;

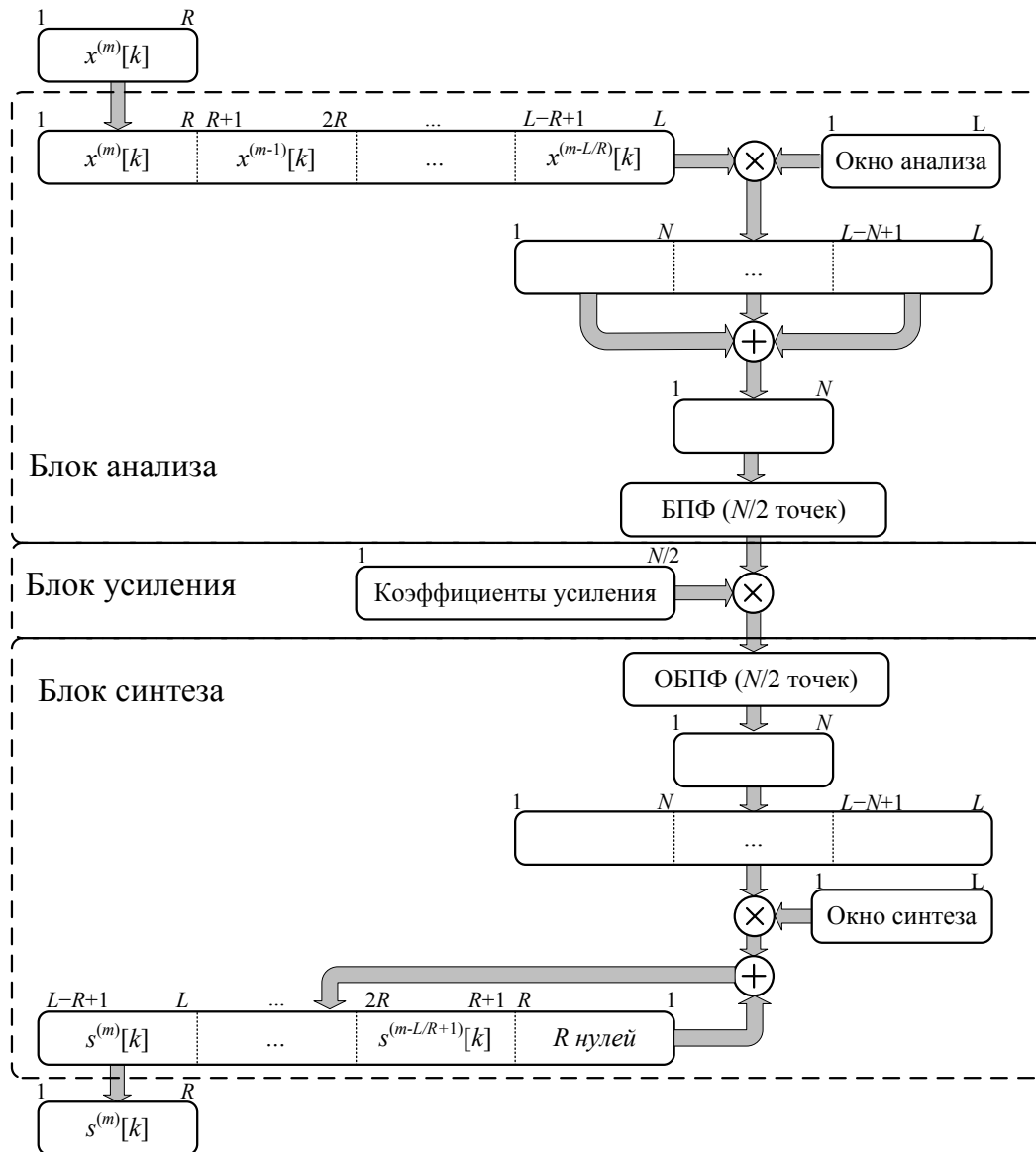


Рис. 1.8. Структура банка фильтров WOLA

- коэффициент передискретизации $\mathcal{O} = N/R = 2$

Работая на указанной частоте дискретизации, банк фильтров имеет 64 частотные полосы (центры полос разнесены на 125 Гц) в диапазоне 0–8 кГц. Можно показать, что задержка банка фильтров WOLA для указанных значений параметров составляет 16 мс, что превышает требуемое время обработки сигнала в современных слуховых аппаратах. Как отмечается в [3], при такой задержке речь собеседника воспринимается хорошо, не возникает рассинхронизации в восприятии речевого сигнала с видимым движением губ. Дискомфорт возникает во время произнесения слов пользователем слухового аппарата, поскольку его собственный голос попадает в барабанную перепонку из нескольких источников: посредством костной

проводимости и после прохождения через слуховой аппарат. Данный эффект субъективно может восприниматься как эхо. Еще одним недостатком указанной подхода является то, что используемый для частотного анализа банк фильтров WOLA не согласован со слуховой системой человека.

Слуховой аппарат с использованием фильтрации во временной области. В [3,4] предлагается подход, использующий фильтрацию во временной области для обработки сигнала в слуховом аппарате. Блок-схема системы, реализующей данный подход, показана на рис. 1.9.

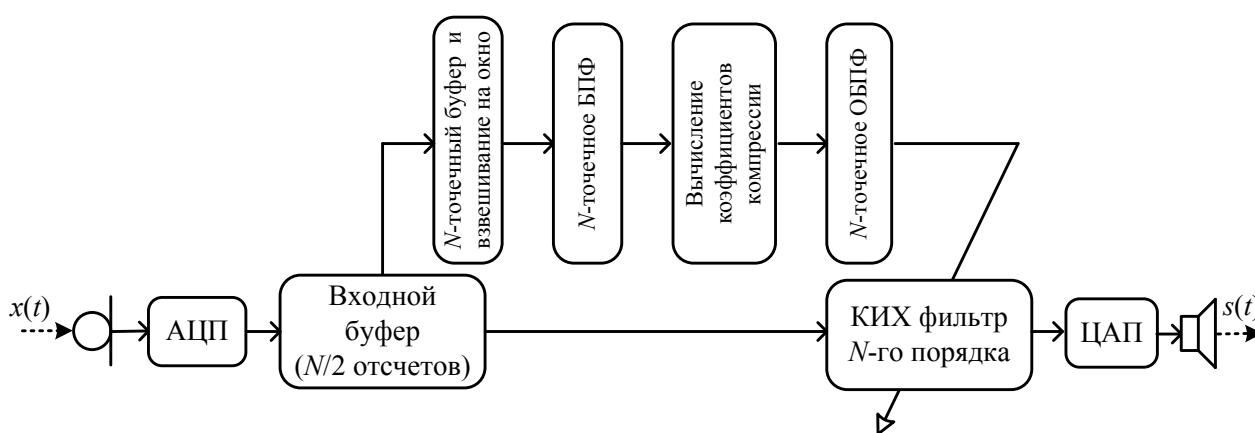


Рис. 1.9. Структура слухового аппарата на основе фильтрации во временной области

Особенностью схемы является малая алгоритмическая задержка, достигаемая за счет того, что анализ сигнала проводится в частотной области, а обработка – во временной. Элементом, выполняющим обработку сигнала, является КИХ-фильтр, коэффициенты которого адаптируются в частотной области. Представленная схема обработки сигнала устраняет один из главных недостатков двух предыдущих подходов: большую алгоритмическую задержку. В [3] предложено применить технику частотного растяжения (*frequency warping*) к схеме на рис. 1.9 для согласования частотного анализа сигнала со слуховой системой человека. Негативным эффектом применения частотного растяжения является увеличение групповой задержки сигнала на низких частотах. Так, если в схеме выбрать $N = 32$, то задержка на низких частотах будет составлять около 6 мс, а на высоких приблизительно порядка 3 мс. Такие значения являются приемлемыми. При этом частотный анализ выполняется в 17 частотных полосах, каждая из которых

имеет ширину приблизительно равную 1,3 барка [45].

1.5.3. Структура слухового аппарата с обработкой в перцептуальной частотной шкале

В [46] для обработки сигнала в прямом канале слухового аппарата предложено использовать неравнополосный косинусно-модулированный банк фильтров с малой групповой задержкой. Данный подход удовлетворяет основным требованиям, предъявляемым к средству частотного анализа в слуховом аппарате. Банк фильтров согласован со слуховой системой человека, поскольку может аппроксимировать различные психоакустические шкалы [47, 48], а также путем выбора соответствующей длины фильтра-прототипа обеспечить малую групповую задержку – менее 4 мс.

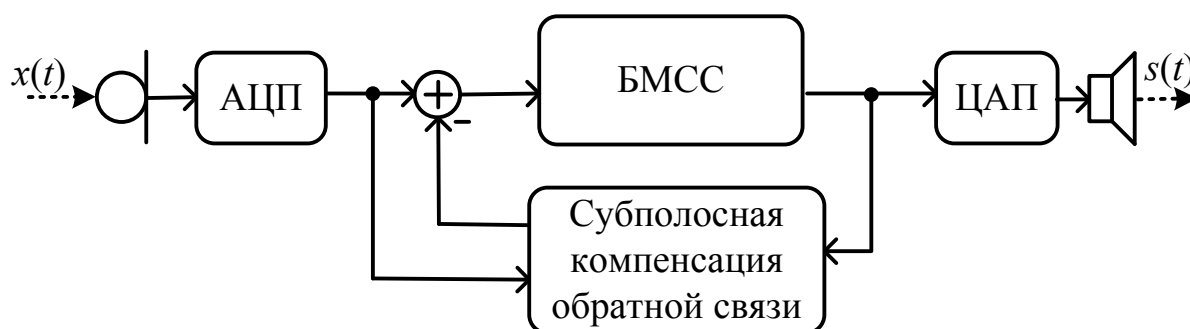


Рис. 1.10. Структура предлагаемого слухового аппарата (БМСС – блок модификации спектра сигнала)

Общая структура предлагаемого слухового аппарата показана на рис. 1.10. Основным компонентом структуры слухового аппарата является блок модификации спектра сигнала (БМСС). БМСС выполняет следующие функции: компрессию динамического диапазона, шумоподавление и компенсацию потери слуха. Другой важной частью структуры является адаптивная система компенсации эффекта АОС, которая не увеличивает задержку сигнала в прямом канале слухового аппарата. В [49] показывается, что адаптивные алгоритмы работают эффективнее, если предварительно входной сигнал разбить на субполосы. Кроме того, последние результаты в области адаптивной фильтрации показывают, что неравнополосные адаптивные структуры в ряде случаев превосходят равнополосные по таким

параметрам, как скорость сходимости и/или модельная ошибка [50]. Поэтому в адаптивной системе компенсации эффекта акустической обратной связи предлагается использовать неравнополосный банк фильтров [51].

Важным вопросом создания процессора слухового аппарата является выбор вычислительной платформы. В этом отношении основными требованиями являются производительность, низкое энергопотребление, низкая тактовая частота и архитектура, позволяющая эффективно выполнять требуемые алгоритмы обработки. В настоящее время существует две основные альтернативы: процессоры цифровой обработки сигналов (ПЦОС) и программируемые логические интегральные схемы (ПЛИС). Преимуществом ПЦОС является их высокая производительность, наличие аппаратных ускорителей операций ЦОС, а также короткие сроки разработки и отладки программного кода. Недостатком ПЦОС для их использования в качестве вычислителей для слуховых аппаратов является тот факт, что для них характерна высокая тактовая частота. Это в свою очередь приводит к повышенному энергопотреблению, что не соответствует требованиям целевого приложения. Необходимо учитывать, что предлагаемые алгоритмы обработки сигнала имеют высокую степень параллелизма, что делает возможным их эффективную реализацию на ПЛИС в виде параллельно работающих структур. Использование параллельных структур позволяет уменьшить тактовую частоту процессора слухового аппарата, поскольку данные в этом случае будут обрабатываться в темпе их поступления.

Реализация предлагаемого подхода к построению слухового аппарата предполагает решение следующих задач.

1. Необходимо разработать алгоритмы, реализующие основные функции слухового аппарата, такие как компрессия динамического диапазона, шумоподавление и компенсация потери слуха. При этом для частотного анализа сигнала должен использоваться неравнополосный косинусно-модулированный банк фильтров.

2. Требуется разработать метод проектирования неравнополосного банка фильтров для аппроксимации психоакустических частотных шкал. Это позволит выполнять в слуховом аппарате обработку сигнала, согласованную с работой слуховой системы человека.

3. Необходимо разработать адаптивный алгоритм компенсации эф-

фекта АОС, который бы не вносил дополнительной задержки в прямой канал слухового аппарата.

4. Для сокращения аппаратных затрат необходимо разработать быстрые алгоритмы реализации неравнополосного банка фильтров, а также других вычислительно-затратных блоков процессора слухового аппарата.

5. В целях уменьшения потребляемой мощности процессора слухового аппарата требуется разработать методы и алгоритмы аппаратной реализации неравнополосных банков фильтров в виде параллельных структур для вычислительных платформ на основе микросхем с программируемой логикой.

1.6. Обзор методов построения неравнополосных банков фильтров

В разделе обсуждаются различные подходы к построению неравнополосных банков фильтров, поскольку таковые необходимы для выполнения обработки сигнала в прямом канале слухового аппарата, а также для реализации субполосной системы компенсации эффекта АОС.

Банк фильтров состоит из двух частей (рис. 1.11): банка фильтров анализа и банка фильтров синтеза [52–56]. Банк фильтров анализа $H_0(z), H_1(z), \dots, H_{M-1}(z)$ выполняет декомпозицию входного сигнала на M полосовых канальных сигналов (субполос). Как правило, для уменьшения нагрузки на вычислительную систему сигналы в каналах банка фильтров подвергаются децимации. В случае, если коэффициенты децимации во всех каналах $S_k = M$, банк фильтров называют *критически децимированным*, иначе банк фильтров считают *передискретизированным*. Синтезирующий банк $F_0(z), F_1(z), \dots, F_{M-1}(z)$ производит интерполяцию канальных сигналов и восстановление исходного сигнала.

Поскольку банк фильтров для слухового аппарата должен быть согласован со слуховой системой человека, которая имеет нелинейную характеристику, то далее рассматриваются подходы к построению неравнополосных банков фильтров.

Существует несколько основных направления при построении нерав-

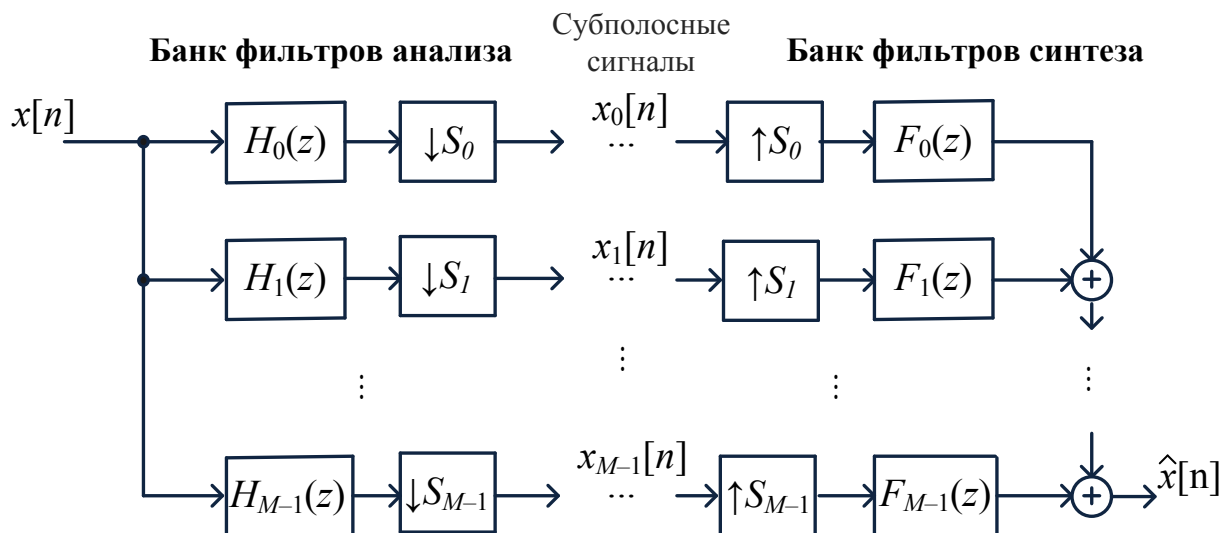


Рис. 1.11. Общая структура банка фильтров

нополосных банков фильтров.

1. Использование пакета дискретного вейвлет-преобразования. В данном подходе для получения неравномерной декомпозиции применяется блок из двух согласованных фильтров нижних частот (НЧ) и высоких частот (ВЧ). Импульсные характеристики данных фильтров представляют собой вейвлет-функции. Вначале исходный сигнал пропускается через блок согласованных фильтров НЧ и ВЧ, после чего выходные сигналы могут быть вновь обработаны при помощи двухканального банка фильтров. Данный процесс продолжается до тех пор, пока не будет получено необходимое частотное разрешение [54, 57]. Пример неравнополосного банка фильтров на основе вейвлет-преобразования приведен на рис. 1.12, где через $H(z)$ и $G(z)$ обозначены согласованные фильтры НЧ и ВЧ соответственно. К недостаткам данного подхода можно отнести нерегулярность дерева фильтров для заданной частотной декомпозиции сигнала, а к достоинствам – хорошо разработанный математический аппарат для эффективной реализации двухканального банков фильтров [58].

2. Неравнополосный банк фильтров можно получить из нескольких секций равнополосных банков фильтров и переходного фильтра [59]. В этом методе переходной фильтр соединяет каналы банка фильтров с различной шириной полос, позволяя получить схожий с равнополосным банк фильтров. При этом необходимо спроектировать несколько равнополосных банков фильтров, из которых выбираются нужные каналы для получения

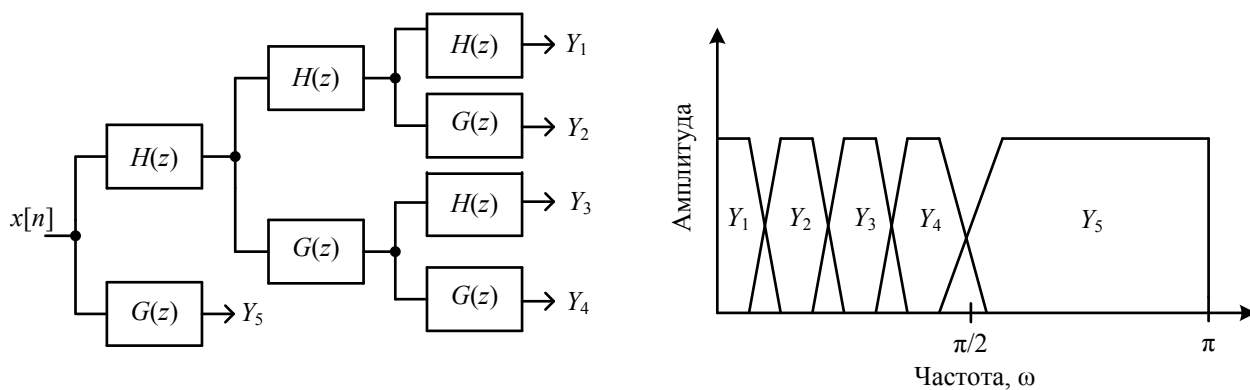


Рис. 1.12. Пример неравнополосного банка фильтров на основе вейвлет-преобразования

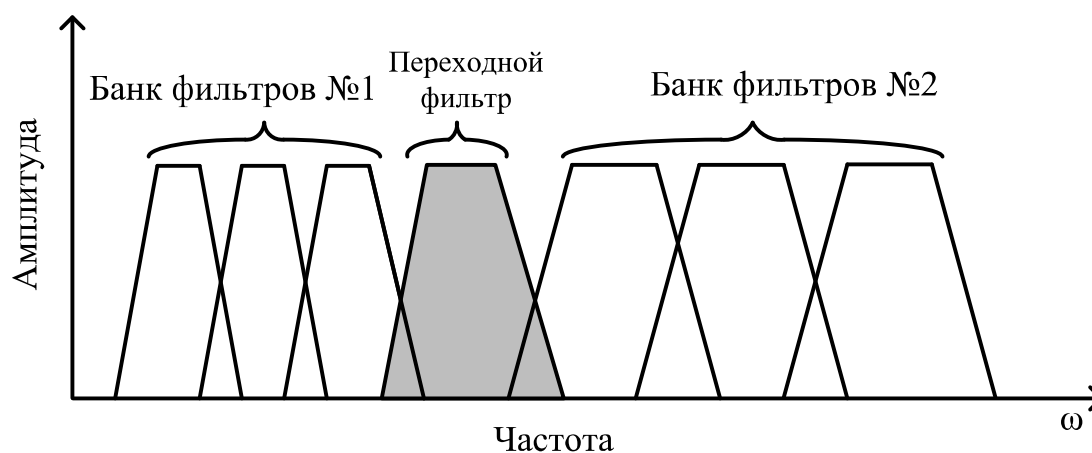


Рис. 1.13. Схема неравнополосный банк фильтров, полученного с использованием переходного фильтра [59]

заданного частотного разделения, а для соединения секций равнополосных банков фильтров специальным образом рассчитывается несимметричный полосовой переходной фильтр (рис. 1.13). Недостатком данного метода является сложность проектирования подобных банков фильтров. Во многих практических задачах число равнополосных банков фильтров, необходимых для получения заданного частотного разбиения, слишком велико. Кроме того поскольку переходной фильтр имеет несимметричную АЧХ, коэффициенты фильтра-прототипа являются комплексными а, следовательно, и сигнал в соответствующем канале также будет комплексным, что создаст неудобства при разработке аппаратной реализации данного типа банков фильтров.

3. Неравнополосный банк фильтров на основе одного равнополосного банка фильтров с объединением каналов [60]. Подход предполагает про-

ектирование равнополосного банка фильтров с большим числом каналов, после чего фильтры анализа объединяются в группы для получения неравнополосного банка фильтров. Данный метод считается эффективным, т.к. для проектирования равнополосных банков фильтров разработано большое число эффективных методов. Недостатком метода является высокая вычислительная сложность при большом количестве каналов.

4. Неравнополосный банк фильтров на основе фазового преобразования (*allpass transform*). В данном методе используется то свойство деформации частотной оси при замене всех элементов задержки системы на фазовые звенья [48, 61]. Таким образом, неравнополосный банк фильтров получается из равнополосного путем деформации частотной оси при помощи фазового преобразования. Такой подход был предложен в [61], где в качестве исходного равнополосного был выбран ДПФ-модулированный банк фильтров. Недостаток такого банка фильтров в том, что выходные сигналы после фильтров анализа являются комплексными, что доставляет дополнительные трудности при аппаратной реализации банка фильтров. Для преодоления этой трудности в [48] было предложено использовать косинусно-модулированный банк фильтров (КМБФ) в качестве исходного. Достоинство данного подхода в том, что все сигналы банка фильтров представляются действительными числами. Кроме того этот метод позволяет изменять частотную характеристику банка анализа и синтеза, варьируя лишь одним коэффициентом фазового преобразования. Недостаток данного метода – необходимость реализации фазового преобразования, которое аппаратно представляет собой каскадно соединенную цепь БИХ-фильтров первого порядка.

В [62] указывается, что банки фильтров на основе фазового преобразования обладают меньшей групповой задержкой и вычислительной сложностью, чем сравнимые с ними неравнополосные банки фильтров с древовидной структурой на основе пакета дискретного вейвлет-преобразования. Поэтому в данной работе рассматривается банк фильтров, получаемый с использованием фазового преобразования. По указанным выше причинам в качестве исходного банка выбирается КМБФ.

1.7. Краткие выводы

Рассмотрение существующих подходов построения слуховых аппаратов показывает, что специфика решаемой ими задачи накладывает значительные ограничения на используемые методы обработки сигнала, а также на структуру процессора слухового аппарата. Показаны недостатки каждого подхода, проявляющиеся либо в значительной задержке сигнала в слуховом аппарате, либо в том, что сигнал обрабатывается в субполосах, не согласованных со слуховой системой человека.

В главе также выполнен обзор и сравнительный анализ методов синтеза неравнополосных банков фильтров для процессоров слуховых аппаратов, на основании которого сделан вывод о необходимости использования банка фильтров на базе фазового преобразования, деформирующего частотную ось.

В качестве альтернативы существующим предложена структура слухового аппарата на основе неравнополосного банка фильтров, выполняющая функции компенсации потери слуха, компрессии динамического диапазона речевого сигнала, шумоподавления и адаптивной компенсации акустической обратной связи.

ГЛАВА 2

СИНТЕЗ НЕРАВНОПОЛОСНОГО КОСИНУСНО-МОДУЛИРОВАННОГО БАНКА ФИЛЬТРОВ

2.1. Косинусно-модулированный банк фильтров (КМБФ)

2.1.1. Основные понятия

Для построения неравнополосного банка фильтров, предложенного в [48], исходным пунктом служит M -канальный КМБФ [63]. Данный банк фильтров образуется путем косинусной модуляции КИХ-фильтра нижних частот с линейной ФЧХ, который называется фильтром-прототипом. Эта идея позволяет свести проектирование целого банка фильтров к расчету одного фильтра-прототипа. Импульсные характеристики фильтров анализа ($h_k[n]$) и синтеза ($f_k[n]$) КМБФ имеют следующий вид:

$$h_k[n] = 2h[n] \cos \left(\frac{(2k+1)\pi}{2M} \left(n - \frac{N-1}{2} \right) + \theta_k \right); \quad (2.1)$$

$$f_k[n] = 2h[n] \cos \left(\frac{(2k+1)\pi}{2M} \left(n - \frac{N-1}{2} \right) - \theta_k \right), \quad (2.2)$$

где N – порядок фильтра-прототипа;

$h[n]$ – коэффициенты фильтра-прототипа нижних частот;

$0 \leq n < N$ – временной индекс;

$0 \leq k < M$ – номер канала.

Схематично АЧХ фильтра-прототипа и всего банка фильтров показаны на рис. 2.1.

Появление банка фильтров, описываемого выражениями (2.1) и (2.2), связано с развитием техники субполосного кодирования [64], где основной решаемой задачей является сокращение объема передаваемой информации. Поэтому предполагалось, что выходные сигналы банка фильтров

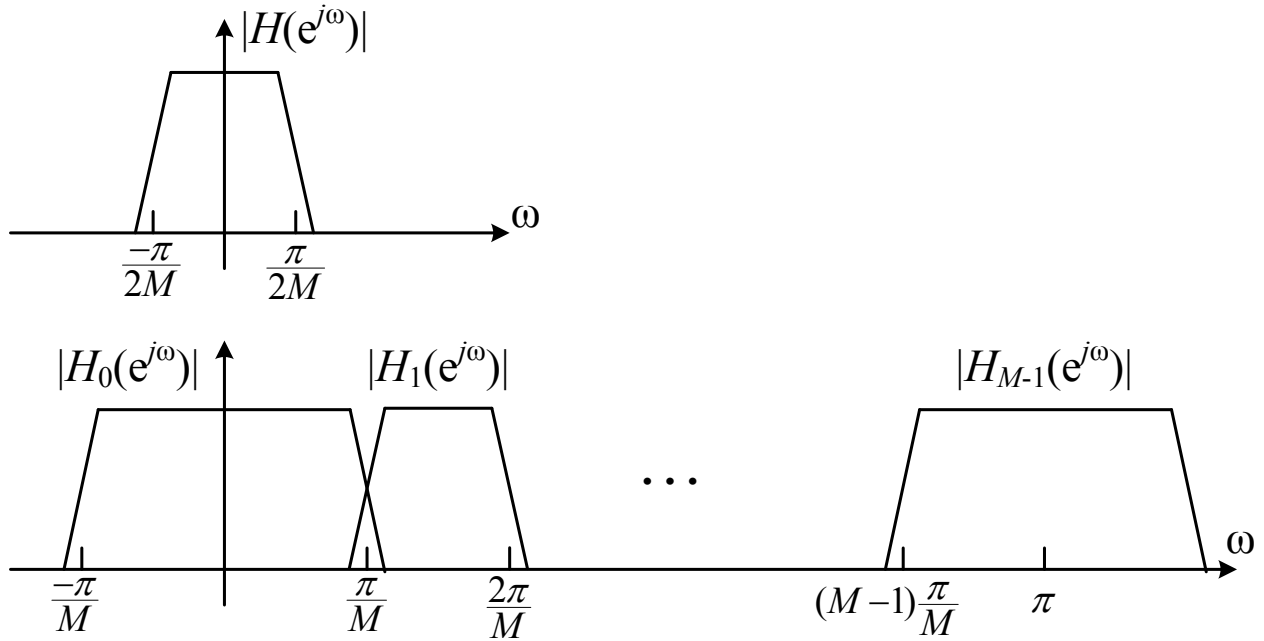


Рис. 2.1. АЧХ фильтра-прототипа и косинусно-модулированного банка фильтров

будут децимированы с коэффициентом M . Данное обстоятельство приводит к появлению компонент наложения спектров значительной величины (*significant aliasing terms*) из соседних полос вследствие неидеальности АЧХ фильтра-прототипа. Наложение спектров из полос, несмежных с данной, определяется ослаблением АЧХ фильтра-прототипа в полосе заграждения. Подавление компонент наложения спектров значительной величины может быть выполнено лишь на этапе реконструкции сигнала с использованием информации из соседних каналов [65].

Рассмотрим механизм подавления компонент наложения спектров значительной величины. Пусть $H(z) = \sum_{n=0}^{N-1} h[n]z^{-n}$ — z -образ фильтра-прототипа. Для $0 \leq k < M$ введем следующие обозначения:

$$U_k(z) = b_k H(zW_{2M}^{k+\frac{1}{2}}); \quad V_k(z) = b_k^* H(zW_{2M}^{-(k+\frac{1}{2})}),$$

где $W_M = e^{-j2\pi/M}$; $b_k = W_{2M}^{\frac{N-1}{2}(k+\frac{1}{2})}$. Тогда для КМБФ анализа и синтеза могут быть записаны выражения:

$$H_k(z) = a_k U_k(z) + a_k^* V_k(z); \quad (2.3-а)$$

$$F_k(z) = a_k^* U_k(z) + a_k V_k(z), \quad (2.3-б)$$

где $a_k = e^{j\theta_k}$. Ключевым моментом при построении КМБФ является предположение, что фильтры несмежных каналов не перекрываются. Например,

в 8-канальном банке фильтров АЧХ четвертого канала $|H_4(e^{j\omega})|$ перекрываются только с $|H_3(e^{j\omega})|$ и $|H_5(e^{j\omega})|$. Полосы пропускания всех остальных фильтров лежат в полосе заграждения фильтра $|H_4(e^{j\omega})|$. Запишем выражение для выхода k -го канала критически децимированного банка фильтров синтеза с учетом компонент наложения спектров значительной величины из смежных каналов:

$$Y_k(z) = \frac{1}{M} F_k(z) \left[a_k U_k(z) X(z) + a_k^* V_k(z) X(z) + \right. \\ \left. + a_k U_k(z W_M^{-k}) X(z W_M^{-k}) + a_k^* V_k(z W_M^k) X(z W_M^k) + \right. \\ \left. + a_k U_k(z W_M^{-(k+1)}) X(z W_M^{-(k+1)}) + a_k^* V_k(z W_M^{k+1}) X(z W_M^{k+1}) \right]. \quad (2.4)$$

В выражение (2.4) входят четыре компоненты наложения спектров, сгруппируем их и введем обозначения:

$$A_{k,\text{low}}(z) = F_k(z) \left[a_k U_k(z W_M^{-k}) X(z W_M^{-k}) + a_k^* V_k(z W_M^k) X(z W_M^k) \right]; \\ A_{k,\text{high}}(z) = F_k(z) \left[a_k U_k(z W_M^{-(k+1)}) X(z W_M^{-(k+1)}) + a_k^* V_k(z W_M^{k+1}) X(z W_M^{k+1}) \right].$$

В общем виде выражение для выхода k -го канала банка фильтров синтеза запишется как

$$Y_k(z) = \frac{1}{M} (F_k(z) H_k(z) X(z) + A_{k,\text{low}}(z) + A_{k,\text{high}}(z)).$$

В работе [65] показывается, что подавление компонент наложения спектров значительной величины достигается при условии

$$A_{k+1,\text{low}}(z) = -A_{k,\text{high}}(z), \quad (2.5)$$

которое после упрощений приводит к уравнению

$$a_k^2 + (a_k^*)^2 = 0, \quad 0 \leq k < M. \quad (2.6)$$

Одним из возможных решений, удовлетворяющих (2.6), является $a_k = e^{j\theta_k}$ при $\theta_k = (-1)^k \pi/4$. Таким образом, выбор параметров модуляции θ_k имеет определяющее значение для подавления компонент наложения спектров значительной величины. Заметим, что последнего удастся достичь за счет использования во всех каналах одинаковых коэффициентов децимации $S_k = M$.

2.1.2. Полифазная реализация КМБФ

Интерес к КМБФ [65–67] объясняется хорошо разработанной теорией алгоритмов их реализации. В [63] предлагается эффективная схема реализации КМБФ на основе полифазного представления фильтра-прототипа. Для случая когда порядок фильтра-прототипа $N = 2mM$, где m – произвольное положительное число, передаточные функции КМБФ анализа представляются в виде

$$\begin{bmatrix} H_0(z) \\ H_1(z) \\ \vdots \\ H_{M-1}(z) \end{bmatrix} = \underbrace{\begin{bmatrix} c_{0,0} & c_{0,1} & \cdots & c_{0,N-1} \\ c_{1,0} & c_{1,1} & \cdots & c_{1,N-1} \\ \vdots & \vdots & \ddots & \vdots \\ c_{M-1,0} & c_{M-1,1} & \cdots & c_{M-1,N-1} \end{bmatrix}}_{M_C} \begin{bmatrix} 1 & & & \\ & z^{-1} & & \\ & & \ddots & \\ & & & z^{-(N-1)} \end{bmatrix} \begin{bmatrix} h[0] \\ h[1] \\ \vdots \\ h[N-1] \end{bmatrix}, \quad (2.7)$$

где $c_{k,n} = \cos\left(\frac{\pi(k+\frac{1}{2})}{M}\left(n - \frac{N-1}{2}\right) + (-1)^k \frac{\pi}{4}\right)$ – компоненты матрицы косинусной модуляции M_C . Из свойства периодичности функции косинус имеем

$$c_{k,(l+2pM)} = (-1)^p c_{k,l},$$

что позволяет записать матрицу косинусной модуляции M_C в блочном виде:

$$M_C = \underbrace{\begin{bmatrix} C & -C & C & -C & \dots \end{bmatrix}}_{m \text{ повторений}}, \quad (2.8)$$

где $C = [c_{k,n}]_{0 \leq k < M, 0 \leq n < 2M}$. Регулярная структура матрицы косинусной модуляции дает возможность представить ее в виде следующего произведения:

$$M_C = C \underbrace{\begin{bmatrix} I_{2M} & -I_{2M} & I_{2M} & -I_{2M} & \dots \end{bmatrix}}_{m \text{ повторений}}, \quad (2.9)$$

где I_n – единичная матрица размера $n \times n$. Умножение на матрицу (2.9) требует в m раз меньше операций умножения, чем на матрицу вида (2.8). Далее подставим выражение для матрицы косинусной модуляции (2.9) в (2.7):

$$\begin{aligned} \begin{bmatrix} H_0(z) \\ H_1(z) \\ \vdots \\ H_{M-1}(z) \end{bmatrix} &= C \underbrace{\begin{bmatrix} I_{2M} & -I_{2M} & I_{2M} & -I_{2M} & \dots \end{bmatrix}}_{m \text{ повторений}} \begin{bmatrix} 1 & & & \\ & z^{-1} & & \\ & & \ddots & \\ & & & z^{-(N-1)} \end{bmatrix} \begin{bmatrix} h[0] \\ h[1] \\ \vdots \\ h[N-1] \end{bmatrix} = \\ &= C \begin{bmatrix} h[0] - h[2M]z^{-2M} + \dots \\ h[1]z^{-1} - h[2M+1]z^{-(2M+1)} + \dots \\ \vdots \\ h[2M-1]z^{-(2M-1)} - h[4M-1]z^{-(4M-1)} + \dots \end{bmatrix} = C \begin{bmatrix} G_0(-z^{2M}) \\ z^{-1}G_1(-z^{2M}) \\ \vdots \\ z^{-(2M-1)}G_{2M-1}(-z^{2M}) \end{bmatrix}, \quad (2.10) \end{aligned}$$

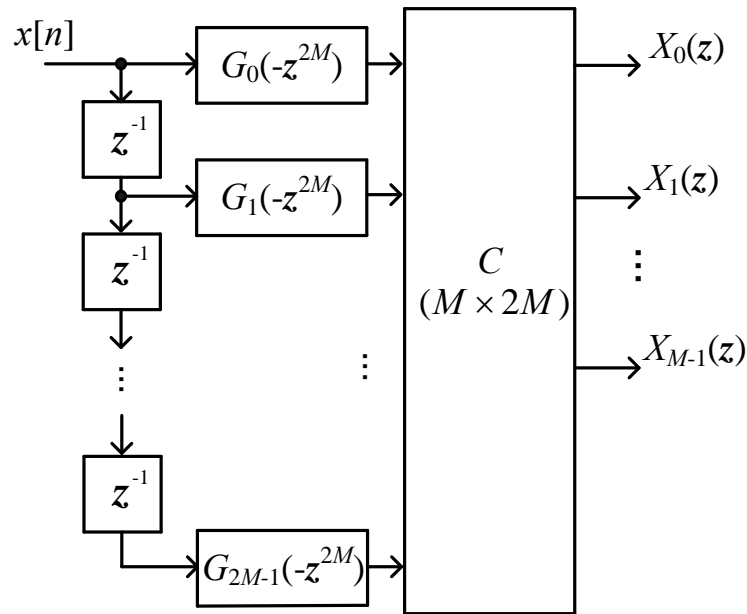


Рис. 2.2. Полифазная реализация косинусно-модулированного банка фильтров анализа

где $G_q(z) = \sum_{p=0}^{m-1} h[q + 2pM]z^p$. Выражение (2.10) служит основой для построения полифазной реализации КМБФ. Элементами вектора-столбца в последнем выражении служат так называемые полифазные компоненты фильтра-прототипа. Аналогичное выражение может быть получено и для банка фильтров синтеза. Эффективная реализация КМБФ анализа на основе выражения (2.10) показана на рис. 2.2. Вычислительная сложность полифазной структуры банка фильтров на рис. 2.2 значительно ниже непосредственной реализации банка в виде гребёнки фильтров. Так, при прямой реализации банка фильтров на один входной отсчёт необходимо выполнить $N \times M$ операций умножения, а при полифазной реализации потребуется $N + M \times 2M$ операций умножения. Преимущество полифазной реализации в том, что умножение входного сигнала на коэффициенты фильтра-прототипа выполняется только один раз, а затем выходные сигналы получают путём косинусной модуляции, для которой при непосредственной реализации требуется $M \times 2M$ операций умножения.

2.1.3. Эффективная реализация косинусной модуляции

Эффективность полифазной реализации КМБФ заключается в возможности построения быстрого алгоритма умножения на матрицу косинусной модуляции. Под быстрым алгоритмом понимается представление матрицы C в виде произведения слабозаполненных матриц. В работе [63] показано, что матрица косинусной модуляции может быть записана в виде

$$C = \begin{cases} \sqrt{M}(-1)^{m_1} \text{DCT-4}_M [(I_M - J_M) - (I_M + J_M)] \\ \text{для чётного } m = 2m_1; \\ \sqrt{M}(-1)^{m_1} \text{DCT-4}_M [(I_M + J_M) (I_M - J_M)] \\ \text{для нечётного } m = 2m_1 + 1, \end{cases} \quad (2.11)$$

где DCT-4_M – матрица дискретного косинусного преобразования (ДКП) 4-го типа, а J_M – диагональная матрицы размера $M \times M$, содержащая единицы на побочной диагонали:

$$\text{DCT-4}_M = \left[\sqrt{\frac{2}{M}} \cos \left(\frac{\pi(k + \frac{1}{2})(\ell + \frac{1}{2})}{M} \right) \right]_{0 \leq k, \ell < M}; \quad J_M = \begin{bmatrix} & & & 1 \\ & & & \\ & & \ddots & \\ & & & \\ 1 & & & \end{bmatrix}.$$

Выражение (2.11) позволяет свести задачу реализации КМБФ к задаче синтеза быстрого алгоритма для ДКП-4. Подход к построению быстрых алгоритмов для всего семейства дискретных синусных и косинусных преобразований, основанный на привлечении методов теории групп и абстрактной алгебры, рассматривается в главе 4.

Рассмотрим вопрос уменьшения количества операций сложения при умножении на матрицу косинусной модуляции. Заметим, что (2.11) представляет собой произведение матрицы предположений $[(I_M - J_M) - (I_M + J_M)]$ на матрицу ДКП-4. При этом для умножения на матрицу предположений необходимо выполнить $3M$ операций сложения. Это число может быть уменьшено до $2M$ [68]. В качестве примера рассмотрим случай, когда m и M – чётные числа, тогда матрица предположений имеет блочный вид:

$$S_M^e = [(I_M - J_M) - (I_M + J_M)] = \begin{bmatrix} I_{M/2} & -J_{M/2} & -I_{M/2} & -J_{M/2} \\ -J_{M/2} & I_{M/2} & -J_{M/2} & -I_{M/2} \end{bmatrix}. \quad (2.12)$$

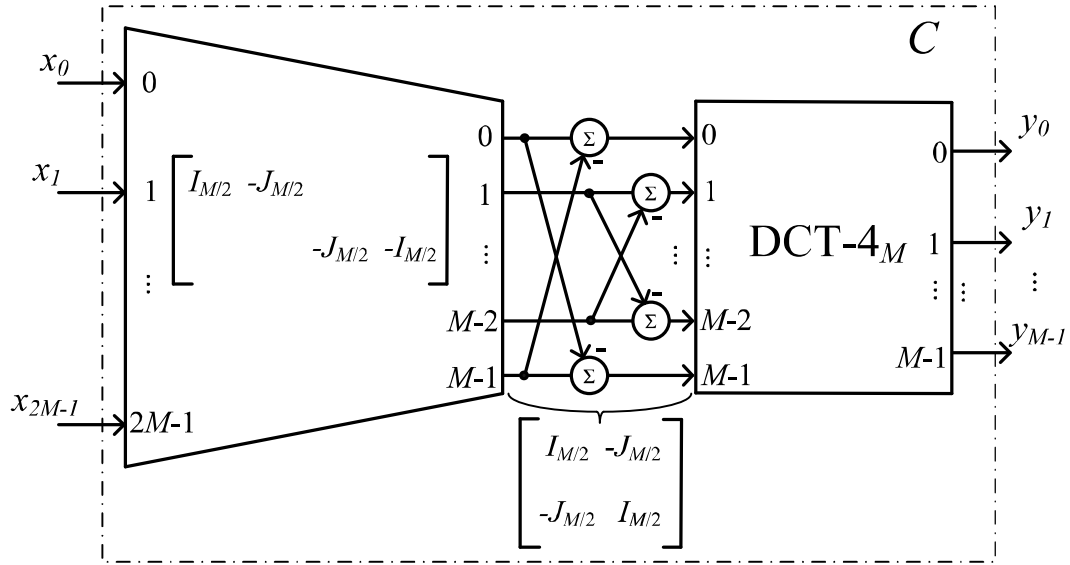


Рис. 2.3. Схема, реализующая умножение на матрицу косинусной модуляции C

Используя элементарные преобразования, приведём (2.12) к блочному ступенчатому виду:

$$\begin{bmatrix} I_{M/2} & \\ J_{M/2} & -I_{M/2} \end{bmatrix} S_M^e = \begin{bmatrix} I_{M/2} & -J_{M/2} & -I_{M/2} & -J_{M/2} \\ & & -2J_{M/2} & -I_{M/2} \end{bmatrix};$$

$$\begin{bmatrix} I_{M/2} & -1/2 J_{M/2} \\ & I_{M/2} \end{bmatrix} \begin{bmatrix} I_{M/2} & \\ J_{M/2} & -I_{M/2} \end{bmatrix} S_M^e = \begin{bmatrix} 1 & \\ & 2 \end{bmatrix} \begin{bmatrix} I_{M/2} & -J_{M/2} & & \\ & & -J_{M/2} & -I_{M/2} \end{bmatrix}.$$

Откуда S_M^e выражается как

$$S_M^e = \begin{bmatrix} I_{M/2} & J_{M/2} \\ -J_{M/2} & I_{M/2} \end{bmatrix} \begin{bmatrix} I_{M/2} & -J_{M/2} & & \\ & & -J_{M/2} & -I_{M/2} \end{bmatrix}. \quad (2.13)$$

Вычисление по формуле (2.13) при реализации требует выполнения $2M$ операций сложения. Аналогичные выражения получаются в случае, если число каналов банка фильтров нечётно [36].

С учётом выражений (2.11) и (2.13) схема, реализующая умножение на матрицу косинусной модуляции C , может быть графически представлена в виде (рис. 2.3).

2.2. Неравнополосный косинусно-модулированный банк фильтров

2.2.1. Фазовое преобразование

При построении неравнополосного банка фильтров ключевую роль играет *фазовое преобразование*, предложенное в [69], где оно использовалось для получения неравномерного частотного разрешения дискретных сигналов. Процедура, описанная в [69], заключается в преобразовании исходной последовательности в новую последовательность, обладающую тем свойством, что её ДПФ равно выборкам z -преобразования исходной последовательности, расположенным через неравные интервалы вокруг единичной окружности. Задача ставилась следующим образом. Пусть $f[n]$ — исходная последовательность, а $g[k]$ — преобразованная последовательность. Необходимо найти линейное преобразование между $f[n]$ и $g[k]$, соответствующее разложению $f[n]$ в систему линейно независимых последовательностей $\psi_k[n]$, которое имеет форму

$$f[n] = \sum_{k=-\infty}^{\infty} g[k]\psi_k[n].$$

От данного преобразования требовалось, чтобы z -преобразования последовательностей $f[n]$ и $g[k]$ были связаны между собой операцией замены переменных, так что

$$G(e^{j\varphi(\omega)}) = F(e^{j\Omega}), \quad (2.14)$$

где $\Omega = \varphi(\omega)$; $G(e^{j\Omega}) = \sum_{k=-\infty}^{\infty} g[k]e^{-j\Omega k}$, а $F(e^{j\omega}) = \sum_{n=-\infty}^{\infty} f[n]e^{-j\omega n}$. Таким образом, функции $\psi_k[n]$ должны иметь характеристику фазового типа, т.е. их z -преобразование на единичной окружности должно иметь единичную амплитуду, независимую от частоты. Если отображение ω на Ω ограничить условием, согласно которому при изменении ω на 2π величина Ω изменяется также на 2π , и потребовать, чтобы z -преобразование $\Psi_k(z)$ от функции $\psi_k[n]$ выражалось рациональными функциями z , то $\Psi_k(z)$ должно иметь форму

$$\Psi_k(z) = \left(\frac{z^{-1} + \alpha}{1 + \alpha z^{-1}} \right)^k. \quad (2.15)$$

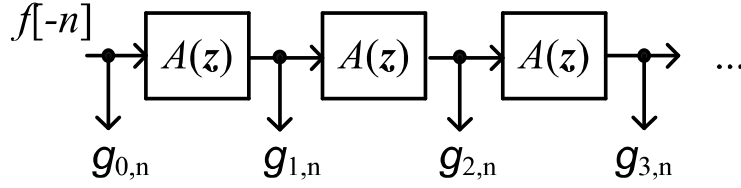


Рис. 2.4. Фазовая цепь, реализующая деформацию частотной оси

Как и требуется, амплитуда $\Psi_k(z)$ для z , расположенных на единичной окружности, равна единице, а функция $\varphi: \omega \rightarrow \Omega$ определяется как

$$\Omega = \varphi(\omega) = \omega - 2 \arctan \frac{\alpha \sin \omega}{1 + \alpha \cos \omega}. \quad (2.16)$$

Можно показать, что обратное соотношение имеет вид

$$\omega = \varphi^{-1}(\Omega) = \Omega - 2 \arctan \frac{\alpha \sin \Omega}{\alpha \cos \Omega - 1}.$$

Поскольку вычисление ДПФ новой последовательности $g[k]$ соответствует равномерной дискретизации по Ω , получающиеся частотные выборки будут соответствовать неравномерной дискретизации по исходной частотной переменной ω . В [69] показывается, что $g[k]$ можно получить, пропуская последовательность $f[-n]$ через линейную инвариантную к сдвигу цепь, показанную на рис. 2.4, где $A(z) = \Psi_1(z)$.

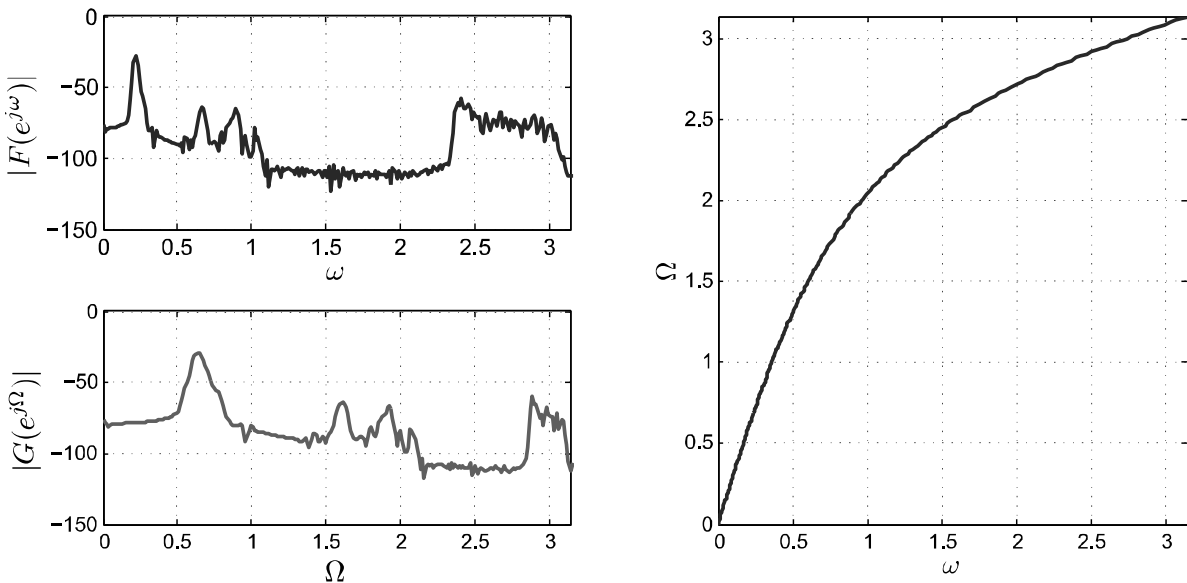


Рис. 2.5. Сигналы $f[n]$ и $g[k]$ в частотной области и график функции деформации частотной оси $\Omega = \varphi(\omega)$

Выходные сигналы цепи $g_{k,n}$ на рис. 2.4 связаны с последовательностью $g[k]$ соотношением $g[k] = g_{k,0}$. Чтобы получить выборки спектра исходной последовательности $f[n]$ на деформированной частотной оси, вычисляют ДПФ последовательности $g[k]$. Пример деформации частотной для спектрального анализа речевого сигнала приведен на рис. 2.5. Параметр деформации устанавливался равным $\alpha = 0,5$, в каждом случае число чисто частотных выборок равно 512.

2.2.2. Деформация частотной оси КМБФ

Неравнополосную версию КМБФ можно получить, используя фазовое преобразование для деформации частотной оси равнополосного КМБФ. Для этого необходимо заменить в выражении (2.10) элементы задержки на фазовые звенья:

$$z^{-1} \rightarrow A(z). \quad (2.17)$$

В рамках данной работы рассматривается применение фазовых звеньев первого порядка с действительными коэффициентами:

$$A(z) = \frac{z^{-1} + \alpha}{1 + \alpha z^{-1}}, \quad \alpha \in \mathbb{R}, \quad |\alpha| < 1, \quad A(e^{j\omega}) = e^{j\varphi(\omega)},$$

где $\varphi(\omega)$ определена в (2.16). В результате замены (2.17) происходит отображение (деформация) оси частот $\omega \mapsto \varphi(\omega)$, что позволяет получить неравнополосный КМБФ (рис. 2.6).

На рис. 2.6 видно, что степень деформации частотной оси зависит только от одного параметра α , регулируя который, можно изменять ширину полос банка фильтров. Полифазная структура неравнополосного КМБФ (рис. 2.7) строится на основании выражений (2.10) и (2.17), где через g_i обозначены коэффициенты полифазных компонент фильтра-прототипа, которые согласно (2.10) связаны с коэффициентами фильтра-прототипа соотношением

$$g_{q+2pM} = (-1)^p h[q + 2pM], \quad 0 \leq q < M, \quad 0 \leq p < m.$$

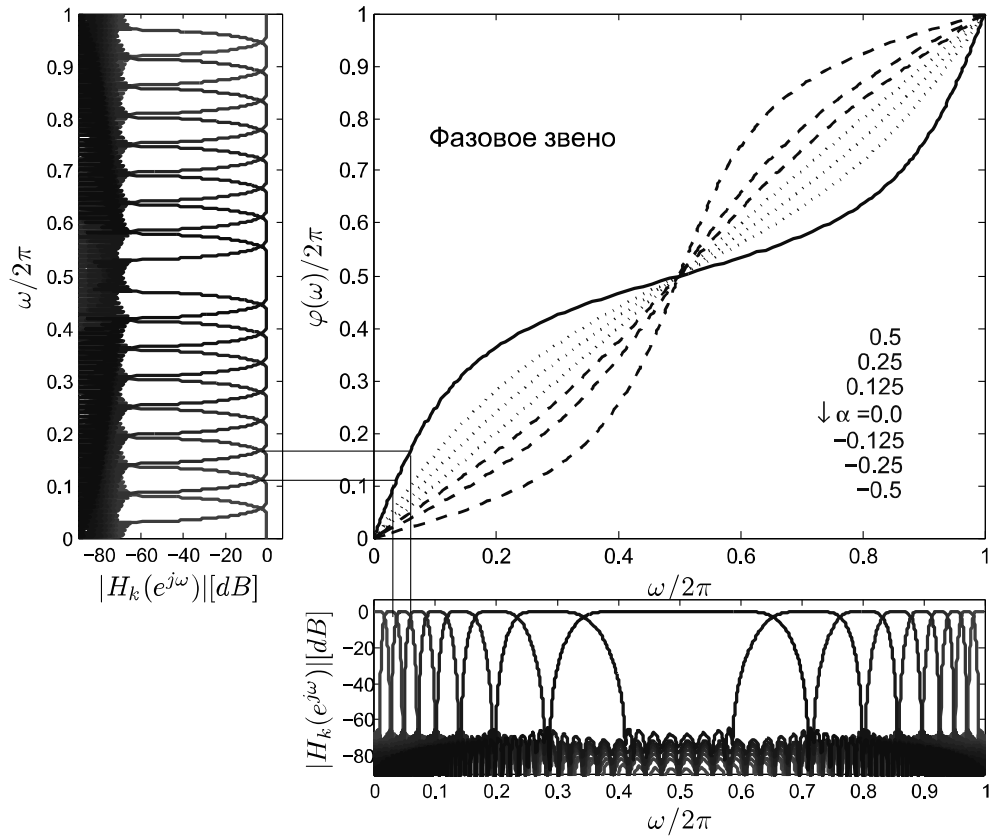


Рис. 2.6. Деформация оси частот при формировании неравнополосного банка фильтров

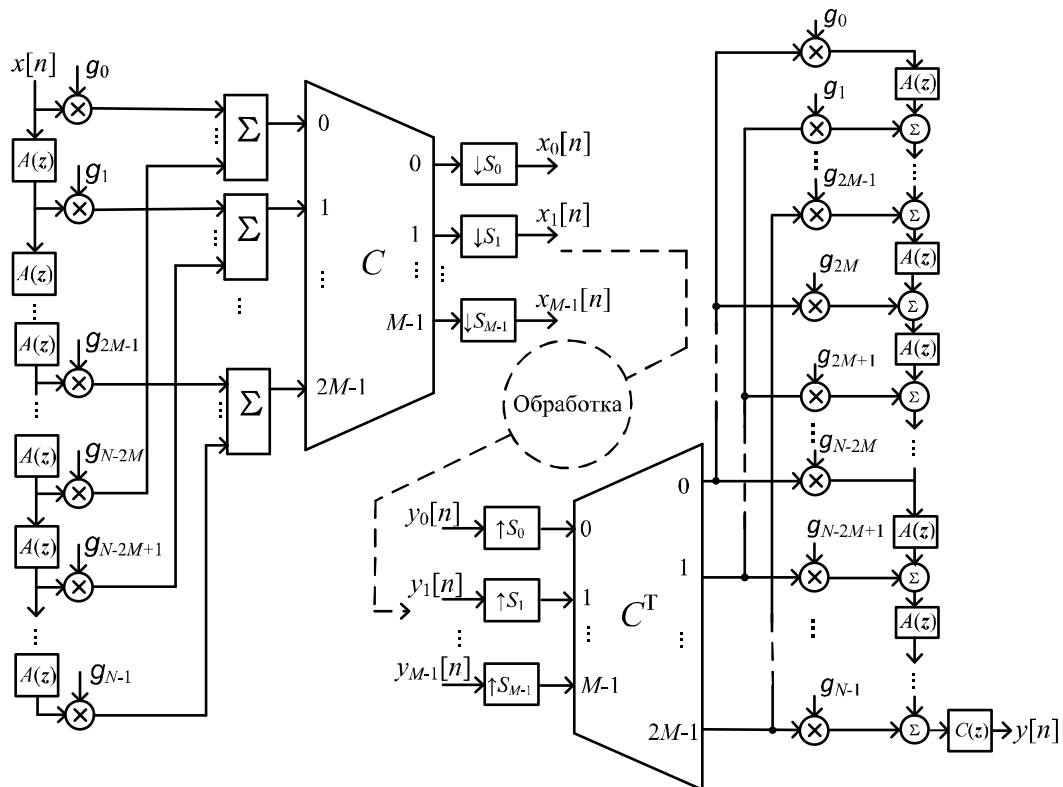


Рис. 2.7. Полифазная структура неравнополосного КМБФ

2.2.3. Расчет коэффициента деформации частотной оси КМБФ для аппроксимации шкалы барков

В слуховых аппаратах, обрабатывающих речевые сигналы и звук, важно иметь неравнополосный банк фильтров, согласованный со слуховой системой человека. Для этого конфигурация полос банка фильтров должна аппроксимировать какую-либо психоакустическую шкалу, в качестве которой чаще всего выбирается шкала барков [45, 47, 48].

В табл. 2.1 приведено соответствие шкалы барков и частотных полос в герцах. Шкала барков отражает тот факт, что частотная избирательность слуховой системы человека намного лучше в области низких частот и ухудшается с увеличением частоты.

В [45] исследовалась возможность отображения шкалы герц в шкалу барков при помощи фазового преобразования. При этом есть всего лишь один параметр, который дает возможность управлять данным отображением, – это коэффициент α в выражении (2.16). Зависимость параметра от частоты дискретизации (f_s) имеет следующий вид [45]:

$$\alpha_{\text{bark}}(f_s) = 0,1957 - 1,048 \left[\frac{\pi}{2} \operatorname{arctg} \left(0,07212 \frac{f_s}{1000} \right) \right]^{1/2} \quad (2.18)$$

Рассмотрим случай, когда частота дискретизации $f_s = 8$ кГц. Данная частота дискретизации соответствует ширине полосы 0–4000 Гц. Поскольку в данный диапазон укладывается 18 критических полос (таблица 2.1), то рассмотрим вопрос построения 18-канального банка фильтров. Подставляя значение $f_s = 8$ кГц в (2.18), получаем значение $\alpha_{\text{bark}} = 0,4092$.

В данной работе предлагается отличный от указанного метод определения параметра фазового преобразования. Для пояснения метода введем выражение для вычисления ошибки аппроксимации k -й критической частотной полосы:

$$E(k) = \frac{|b_{k-1} - c_{k-1}| + |b_k - c_k|}{|b_k - b_{k-1}|}, \quad (2.19)$$

где b_{k-1} и b_k – границы k -й критической частотной полосы (в герцах), а c_{k-1} и c_k – это значения границ полос пропускания банка фильтров, получаемого при заданном значении параметра α_{bark} . Выражение (2.19) позволяет

Таблица 2.1. Шкала барков

№ критической полосы	Частотный диапазон, Гц	№ критической полосы	Частотный диапазон, Гц
1	0–100	13	1720–2000
2	100–200	14	2000–2320
3	200–300	15	2320–2700
4	300–400	16	2700–3150
5	400–510	17	3150–3700
6	510–630	18	3700–4400
7	630–770	19	4400–5300
8	770–920	20	5300–6400
9	920–1080	21	6400–7700
10	1080–1270	22	7700–9500
11	1270–1480	23	9500–12000
12	1480–1720	24	12000–15500

оценить относительную погрешность аппроксимации k -й критической частотной полосы. В (2.19) $|b_{k-1} - c_{k-1}|$ – ошибка аппроксимации нижней границы k -й полосы, а $|b_k - c_k|$ – погрешность аппроксимации верхней границы k -й полосы. Эти ошибки суммируются, и вычисляется их отношение к ширине k -й полосы. Для $\alpha_{\text{bark}} = 0,4092$ график ошибки приведен на рис. 2.8.

Можно сформулировать два критерия минимизации ошибки аппроксимации шкалы барков (2.19): в чебышевском смысле

$$\phi_C = \max_{k \in [1, M]} (E(k))$$

и в среднеквадратичном смысле

$$\phi_S = \frac{1}{M} \sum_{k=1}^M (E(k) - \bar{E}), \quad \bar{E} = \frac{1}{M} \sum_{k=1}^M E(k),$$

где M – количество каналов банка фильтров.

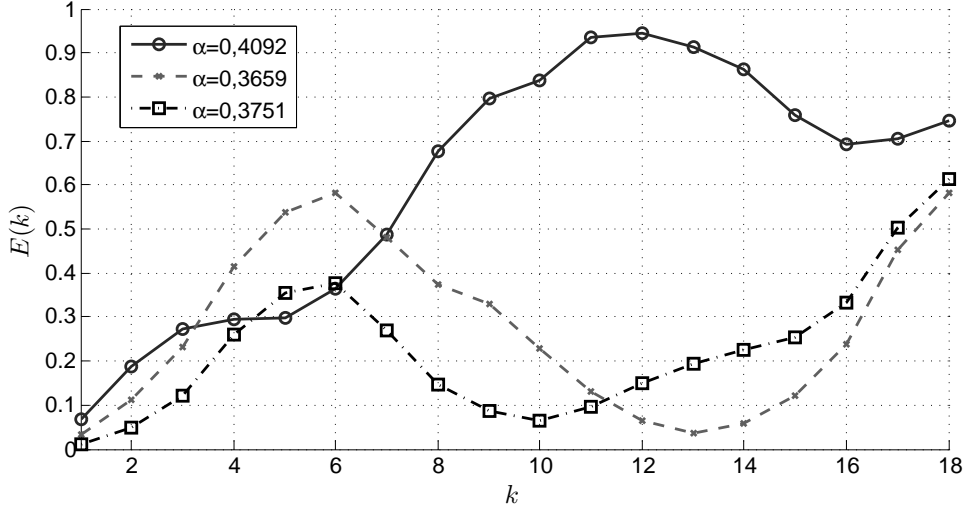


Рис. 2.8. Ошибка аппроксимации шкалы барков при различных значениях α

Единственный параметр, который влияет на значение ϕ_C и ϕ_S , это коэффициент деформации частотной оси α . Таким образом, задачу нахождения оптимального коэффициента α_{opt} можно сформулировать следующим образом:

$$\alpha_{opt}^C = \arg \min_{\alpha} \phi_C \quad (2.20)$$

$$\alpha_{opt}^S = \arg \min_{\alpha} \phi_S \quad (2.21)$$

Для нахождения оптимального значения коэффициента $\alpha_{opt}^{\{C,S\}} \in [0, 1]$ может быть использован любой метод одномерного поиска (метод чисел Фибоначчи, метод золотого сечения и др.). Для $M = 18$ и $f_s = 8000$ кГц минимальное значение функции $\alpha_{opt}^C = 0,3659$, а для $\alpha_{opt}^S = 0,3751$. Графики ошибок, которые получаются в обоих случаях, показаны на рис. 2.8. Приведенные графики показывают, что, используя предлагаемый метод определения параметра деформации частотной оси, можно значительно снизить ошибку аппроксимации банком фильтров шкалы барков.

2.2.4. Выбор коэффициентов децимации/интерполяции

Важным вопросом при проектировании неравнополосного банка фильтров является выбор коэффициентов децимации/интерполяции [70]. Поскольку выходные сигналы банка фильтров анализа являются полосовыми (рис. 2.9), для выбора коэффициента децимации можно воспользоваться

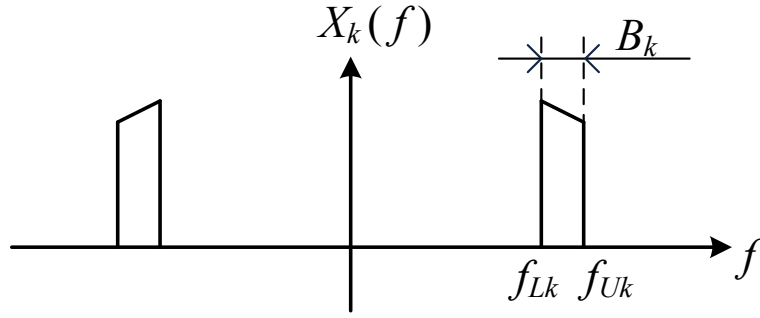


Рис. 2.9. Полосовой сигнал на выходе k -го канала банка фильтров анализа

теорией дискретизации полосовых сигналов [71]. Систематическое описание метода выбора коэффициентов децимации применительно к неравнополосным банкам фильтров дано в [72], где показано, что значения коэффициентов децимации/интерполяции (S_k) должны удовлетворять неравенствам

$$\left\lfloor \frac{n_k}{2f_{U_k}} \right\rfloor \geq S_k \geq \left\lceil \frac{n_k - 1}{2f_{L_k}} \right\rceil, \quad 1 \leq n_k \leq \left\lfloor \frac{f_{U_k}}{f_{U_k} - f_{L_k}} \right\rfloor, \quad (2.22)$$

в которых полагается, что сигнал k -го канала является действительным, занимает частотный диапазон $[-f_{U_k}; -f_{L_k}] \cup [f_{L_k}; f_{U_k}]$ и имеет частоту дискретизации f_s . Оператор $\lfloor a \rfloor$ обозначает наибольшее целое число, меньшее a , а оператор $\lceil a \rceil$ обозначает наименьшее целое число, большее a .

Рассмотрим вопрос определения границ частотных полос f_{U_k} и f_{L_k} . В случае равнополосного банка фильтров можно считать, что диапазон $[f_{L_k}; f_{U_k}]$ совпадает с полосой пропускания k -го фильтра. Данное предположение при использовании механизма подавления компонент наложения спектров значительной величины позволяет получить критически децимированный банк фильтров. Для неравнополосного банка фильтров, полученного путем деформации оси частот (2.16), механизм подавления компонент наложения спектров значительной величины не работает, поскольку смежные полосы имеют различные коэффициенты децимации (условие (2.5) не выполняется). Поэтому при определении f_{L_k} и f_{U_k} необходимо исходить из того же принципа, который лежит в основе механизма подавления компонент наложения спектров значительной величины: *фильтры несмежных каналов не перекрываются*. Это означает, что в диапазон $[f_{L_k}; f_{U_k}]$ должны входить полоса пропускания фильтра k -го канала, а также полосы пропускания $(k - 1)$ -го и $(k + 1)$ -го фильтров.

Определим общий алгоритм нахождения границ частотных полос f_{L_k}

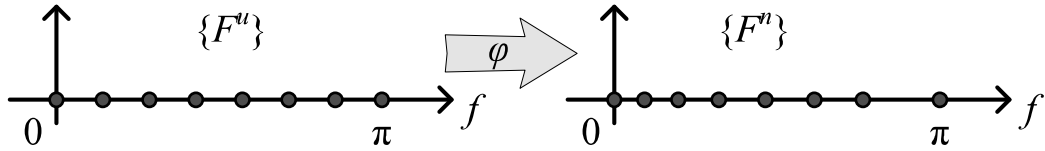


Рис. 2.10. Отображение точек на частотной оси под действием фазового преобразования

и f_{U_k} для неравнополосного КМБФ. Упорядоченное множество

$$F^u = \left\{ \frac{\pi k}{M} \right\}, \quad k = 0, 1 \dots M,$$

представляет собой границы частотных полос пропускания равнополосного КМБФ. Отображение (2.16) позволяет перейти к границам частотных полос неравнополосного КМБФ (рис. 2.10):

$$\varphi: F^u \rightarrow F^n.$$

Тогда согласно предлагаемому правилу частотный диапазон для k -го канала неравнополосного КМБФ определяется как

$$f_{L_k} = \varphi(F_{k-1}^u)/2\pi; \quad f_{U_k} = \varphi(F_{k+2}^u)/2\pi, \quad (2.23)$$

для $k = 0$ и $K = M - 1$ справедливы следующие соотношения:

$$f_{L_0} = 0; \quad f_{U_0} = \varphi(F_2^u)/2\pi; \quad (2.24-a)$$

$$f_{L_{M-1}} = \varphi(F_{M-2}^u)/2\pi; \quad f_{U_{M-1}} = \varphi(F_M^u)/2\pi. \quad (2.24-б)$$

Совместное использование правил определения границ частотных полос (2.23), (2.24) и выражений (2.22) позволяет определить значения коэффициентов децимации/интерполяции S_k для неравнополосного КМБФ.

2.2.5. Выравнивание групповой задержки

Важной проблемой использования банков фильтров в слуховых аппаратах, работающих в реальном времени, является то, что субполосная система анализ/синтез вносит задержку. Растяжение частотной оси приводит к тому, что групповая задержка становится нелинейной функцией,

поскольку банк фильтров на основе фазового преобразования всегда является БИХ-системой. Групповая задержка зависит от силы деформации частотной оси и является монотонной функцией с направлением, определяемым знаком коэффициента фазового преобразования α . Для $\alpha > 0$ нижние частоты задерживаются на число отсчетов большее порядка фильтра, в то время как высокочастотные компоненты задерживаются меньше. Для $\alpha < 0$ ситуация изменяется в противоположную сторону. Заметим также, что отклик системы анализ/синтез с деформированной осью частот эквивалентен отклику цепочки фазовых звеньев из N секций [73].

На рис. 2.11 показаны групповые задержки банков фильтров с различными фильтрами прототипами (порядка N), также с различающимися коэффициентами фазового преобразования (α). Кружками на графиках обозначены точки, в которых групповая задержка такая же, как у соответствующих равнополосных банков фильтров (с таким же фильтром-прототипом). Легко заметить, что отличие между этими значениями является существенным, поэтому остро стоит вопрос выравнивания групповой задержки в системах на основе деформации оси частот.

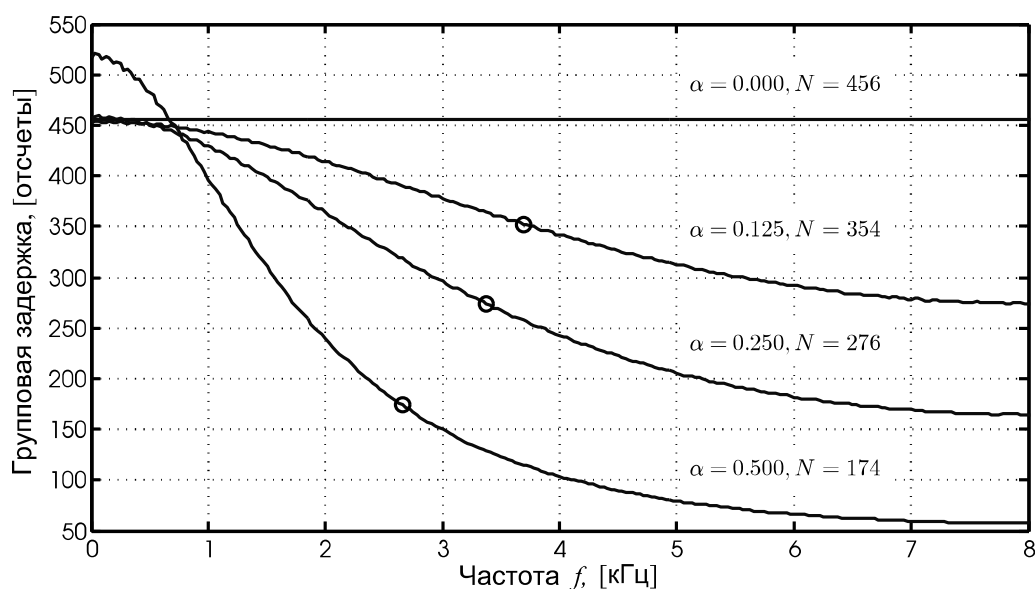


Рис. 2.11. График групповой задержки для различных банков фильтров

Как известно, групповая задержка линейной системы $\tau(\omega)$ равна производной от ФЧХ системы с обратным знаком [42]:

$$\tau(\omega) = -\frac{d\varphi(\omega)}{\omega},$$

из этого следует, что изменение групповой задержки в зависимости от частоты (рис. 2.11) это следствие нелинейности ФЧХ системы анализ/синтез НКМБФ. Нелинейная ФЧХ системы нежелательна, поскольку банк фильтров предполагается использовать для обработки звуковых сигналов. Ниже будет рассмотрен вопрос построения фильтра для коррекции ФЧХ системы анализ/синтез НКМБФ.

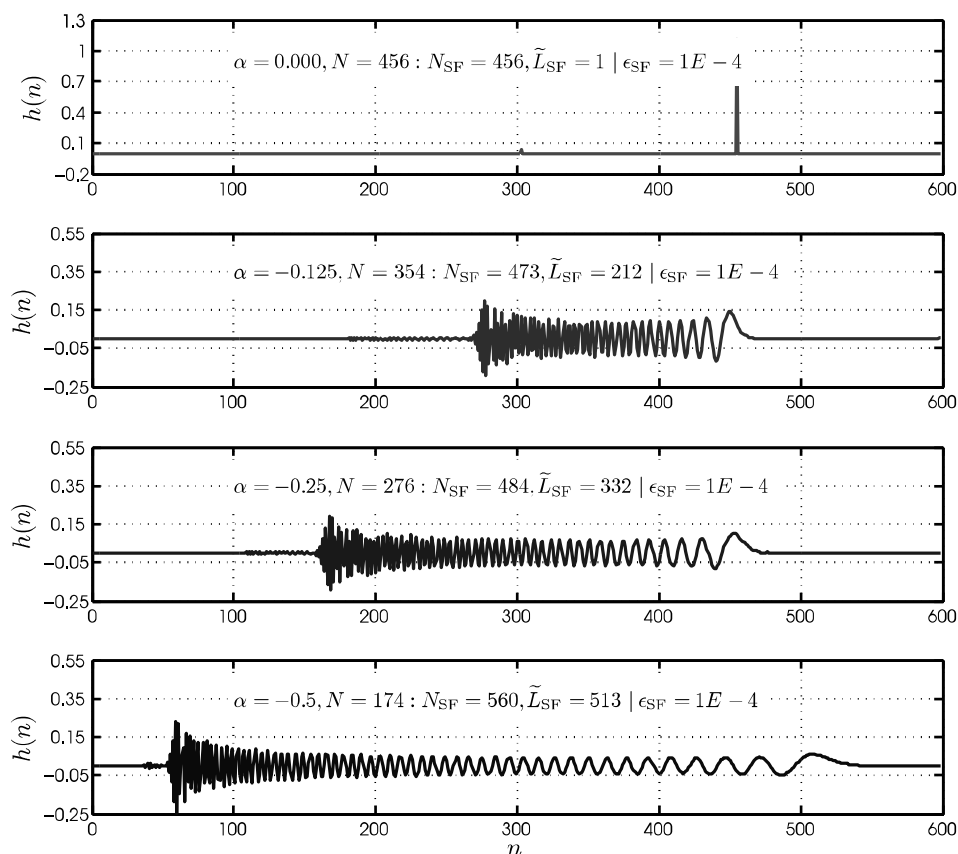


Рис. 2.12. Импульсные характеристики систем анализ/синтез с различными параметрами N и α

В [74] показано, что КИХ-фильтр для выравнивания групповой задержки можно получить путем реверсии импульсной характеристики системы анализ/синтез банка фильтров, для которой требуется корректировка (на рис. 2.7 данный фильтр обозначен через $C(z)$). Даже в случае бесконечного отклика в нем есть определяющий значимый фрагмент, что показано на рис. 2.12. Следовательно, можно принять, что существуют только отсчеты с индексами между 0 и N_{SF} , которые определяют передаточную

функцию полиномиального типа, почти эквивалентную исходной:

$$A^N(z) \approx \sum_{n=0}^{N_{SF}} a[n]z^{-n}.$$

Реверсируя порядок следования отсчетов, получаем КИХ-фильтр, который аппроксимирует ряд соединенных обратных передаточных функций и задержек:

$$C(z) = \sum_{n=0}^{N_{SF}} a[N_{SF} - n]z^{-n} \approx A^N(z^{-1})z^{-N_{SF}}.$$

Поскольку $A^N(z^{-1}) = A^{-N}(z)$, то

$$A^N(z)C(z) \approx A^N(z)A^{-N}(z)z^{-N_{SF}} = z^{-N_{SF}}.$$

Очевидно, что отклонение произведения стоящего слева от чистой задержки зависит от остатка, который остается после усечения отклика. Экспериментально было установлено, что достаточно определить последний значащий отсчет, после которого абсолютные значения становятся меньше порога $\epsilon_{SF} = 10^{-4}$. Для широких диапазонов α и N колебания групповой задержки после корректировки с использованием этого порогового значения не превышают нескольких отсчетов для низких частот и быстро уменьшаются с увеличением частоты.

2.3. Метод расчета фильтра-прототипа неравнополосного косинусно-модулированного банка фильтров

2.3.1. Постановка задачи

Рассмотрим задачу расчета фильтра-прототипа НКМБФ [75]. Для этого запишем его передаточную функцию (см. рис. 2.7)

$$T_{\text{all}}(z) = \sum_{k=0}^{M-1} \sum_{l=0}^{S_k-1} H_k(zW_{S_k}^l)F_k(z). \quad (2.25)$$

Данную функцию можно представить в виде суммы передаточной функции искажения банка фильтров $T_{\text{dist}}(z)$ и передаточной функции компонент наложения спектра $T_{\text{alias}}(z)$:

$$T_{\text{all}}(z) = T_{\text{dist}}(z) + T_{\text{alias}}(z), \quad (2.26)$$

где $T_{\text{dist}}(z) = \sum_{k=0}^{M-1} H_k(z)F_k(z)$; $T_{\text{alias}}(z) = \sum_{k=0}^{M-1} \sum_{l=1}^{S_k-1} H_k(zW_{S_k}^l)F_k(z)$. Условием полного восстановления сигнала в банке фильтров являются равенства

$$T_{\text{dist}}(z) = z^{-D}; \quad T_{\text{alias}}(z) = 0.$$

Во многих случаях этим условиям сложно удовлетворить, поэтому их часто ослабляют, требуя лишь приблизительного равенства. В последнем случае банк фильтров относят к классу систем с квазиполным восстановлением (*near-perfect reconstruction*). К этому классу относится, в частности, и НКМБФ, рассматриваемый в данной работе.

Имеется два источника искажений в реконструированном сигнале в банке фильтров:

1) искажения появляются вследствие неравномерности в полосе пропускания фильтра-прототипа $H(e^{j\omega})$. Этот тип искажений описывается частотной характеристикой искажений $T_{\text{dist}}(e^{j\omega})$;

2) искажения возникают в результате неполного подавления компонент наложения спектров, которые образуются в результате децимации/интерполяции сигналов в каналах банка фильтров. Этот тип искажений описывается частотной характеристикой передачи компонент наложения спектров $T_{\text{alias}}(e^{j\omega})$.

Таким образом, для уменьшения уровня искажений в реконструированном сигнале фильтр-прототип должен быть рассчитан таким образом, чтобы иметь малую неравномерность в полосе пропускания (тем самым минимизируется $T_{\text{dist}}(e^{j\omega})$), а также обеспечивать подавление компонент наложения спектра (для минимизации $T_{\text{alias}}(e^{j\omega})$).

2.3.2. Расчетные выражения

Решим вспомогательную задачу получения основных расчетных выражений. Для дальнейшего изложения необходимо выразить зависимость

функции $T_{\text{all}}(e^{j\omega})$ от коэффициентов фильтров-прототипов $h[n]$.

Известно, что частотную характеристику КИХ-фильтра с линейной ФЧХ можно записать в виде [42]

$$H(\omega) = e^{-j[(N-1)\omega/2]} \mathbf{C}^T(\omega) \mathbf{h},$$

где

$$\mathbf{C}(\omega) = \begin{cases} \left[2 \cos\left(\frac{\omega}{2}\right) \ 2 \cos\left(\frac{3\omega}{2}\right) \ \dots \ 2 \cos\left(\frac{[N-1]\omega}{2}\right) \right]^T, \\ N \text{ четное;} \\ \left[1 \ 2 \cos(\omega) \ \dots \ 2 \cos\left(\frac{[N-1]\omega}{2}\right) \right]^T, \\ N \text{ нечетное;} \end{cases}$$

$$\mathbf{h} = \left[h\left(\left\lfloor \frac{N}{2} \right\rfloor\right) \ h\left(\left\lfloor \frac{N}{2} \right\rfloor + 1\right) \ \dots \ h(N-1) \right]^T.$$

Надстрочный знак T обозначает транспонирование.

Рассмотрим функцию $T_{\text{all}}(e^{j\omega})$ неравнополосного КМБФ, которую с помощью выражений (2.3) и (2.25) можно записать как [76]

$$T_{\text{all}}(e^{j\omega}) = \mathbf{h}^T \mathbf{U}(\omega) \mathbf{h}, \quad (2.27)$$

где

$$\begin{aligned} \mathbf{U}(\omega) = & \sum_{k=0}^{M-1} \sum_{l=0}^{S_k-1} \left[a_k b_k e^{-j[(N-1)\gamma_1(\omega, l, k)/2]} \mathbf{C}^T(\gamma_1(\omega, l, k)) + \right. \\ & \left. + a_k^* b_k^* e^{-j[(N-1)\gamma_2(\omega, l, k)/2]} \mathbf{C}^T(\gamma_2(\omega, l, k)) \right] \times \\ & \times \left[a_k^* b_k e^{-j[(N-1)\gamma_1(\omega, 0, k)/2]} \mathbf{C}^T(\gamma_1(\omega, 0, k)) + \right. \\ & \left. + a_k b_k^* e^{-j[(N-1)\gamma_2(\omega, 0, k)/2]} \mathbf{C}^T(\gamma_2(\omega, l, k)) \right]^T, \end{aligned}$$

и

$$\gamma_1(\omega, l, k) = -\varphi(\omega + 2\pi l/S_k) - 2\pi(k + 0.5)/2M;$$

$$\gamma_2(\omega, l, k) = -\varphi(\omega + 2\pi l/S_k) + 2\pi(k + 0.5)/2M.$$

Таким образом, получено выражение, которое определяет зависимость функции $T_{\text{all}}(e^{j\omega})$ от коэффициентов фильтра-прототипа НКМБФ.

2.3.3. Расчёт фильтра-прототипа НКМБФ

Рассмотрим задачу нахождения коэффициентов фильтра-прототипа банка фильтров, которые обеспечивают минимальный уровень вносимых искажений в реконструированный сигнал неравнополосного КМБФ. Для этого сформулируем задачу оптимизации:

$$\min_{\mathbf{h}} \max |E(\omega)|, \quad (2.28)$$

где $E(\omega)$ — функция ошибки (мера искажений, вносимых банком фильтров), определяемая как

$$E(\omega) = |T_{\text{all}}(e^{j\omega})|^2 - 1, \quad \omega \in [0, \pi],$$

а \mathbf{h} — вектор коэффициентов фильтра-прототипа.

Использование выражения (2.27) для передаточной функции банка фильтра позволяет записать функцию ошибки в виде [77]

$$E(\omega) = (\mathbf{h}^T \mathbf{U}_r(\omega) \mathbf{h})^2 + (\mathbf{h}^T \mathbf{U}_i(\omega) \mathbf{h})^2 - 1,$$

где

$$\mathbf{U}_r(\omega) = \text{Re}(\mathbf{U}(\omega)); \quad \mathbf{U}_i(\omega) = \text{Im}(\mathbf{U}(\omega)).$$

Решения минимаксной задачи (2.28) можно достичь, минимизировав взвешенную по наименьшим квадратам функцию ошибки [66]. Для этого сформулируем целевую функцию

$$g = \sum_{\omega \in [0, \pi]} B(\omega) |E(\omega)|^2,$$

где $B(\omega)$ — неотрицательная весовая функция.

Для проектирования предлагается использовать подход [66], состоящий из двух итеративных процедур, одна из которых располагается внутри другой. Индекс ℓ будет использоваться для значений функций на ℓ -й итерации внутреннего цикла оптимизации, а индекс μ — для обозначения величин, которые изменяются на μ -й итерации внешнего цикла. Внутренний цикл итерационно оптимизирует коэффициенты фильтра-прототипа (\mathbf{h}_ℓ), в то время как во внешнем цикле процедуры итеративно корректируется

значение весовой функции $B_\mu(\omega)$. Учитывая данные замечания, целевая функция на ℓ -й итерации принимает следующий вид:

$$g(\mathbf{h}_\ell) = \sum_{\omega \in [0, \pi]} B(\omega) \left[(\mathbf{h}_\ell^T \mathbf{U}_r(\omega) \mathbf{h}_\ell)^2 + (\mathbf{h}_\ell^T \mathbf{U}_i(\omega) \mathbf{h}_\ell)^2 - 1 \right]^2.$$

Используя введённые обозначения, найдём градиент целевой функции. Для этого вначале определим градиент функции ошибки:

$$\nabla E_\ell(\omega) = \nabla E_r + \nabla E_i,$$

где

$$\nabla E_r = 2\mathbf{h}_\ell \mathbf{U}_r(\omega) \mathbf{h}_\ell (\mathbf{U}_r(\omega) + \mathbf{U}_r^T(\omega)) \mathbf{h}_\ell;$$

$$\nabla E_i = 2\mathbf{h}_\ell \mathbf{U}_i(\omega) \mathbf{h}_\ell (\mathbf{U}_i(\omega) + \mathbf{U}_i^T(\omega)) \mathbf{h}_\ell.$$

Учитывая данные выражения, градиент целевой функции принимает вид

$$\nabla g(\mathbf{h}_\ell) = 2 \sum_{\omega \in [0, \pi]} B_\mu(\omega) E_\ell(\omega) \cdot \nabla E_\ell(\omega).$$

Предположим, что $\mathbf{h}_{\ell, opt} = \mathbf{h}_\ell + \mathbf{e}_\ell$ является оптимальным решением, которое обращает градиент целевой функции в нуль. Для определения вектора невязки \mathbf{e}_ℓ необходимо разложить градиент $\nabla g(\mathbf{h}_\ell)$ в ряд Тейлора в окрестности точки оптимума:

$$\nabla g(\mathbf{h}_\ell + \mathbf{e}_\ell) = \nabla g(\mathbf{h}_\ell) + \nabla^2 g(\mathbf{h}_\ell) \cdot \mathbf{e}_\ell + \dots = 0.$$

Не учитывая члены высших порядков, можно получить выражение для вектора невязки

$$\nabla g(\mathbf{h}_\ell) = -\nabla^2 g(\mathbf{h}_\ell) \mathbf{e}_\ell,$$

откуда

$$\mathbf{e}_\ell = -[\nabla^2 g(\mathbf{h}_\ell)]^{-1} \cdot \nabla g(\mathbf{h}_\ell), \quad (2.29)$$

где $\nabla^2 g(\mathbf{h}_\ell)$ — матрица Гессе целевой функции, которая определяется как

$$\begin{aligned} \nabla^2 g(\mathbf{h}_\ell) = & 2 \sum_{\omega \in [0, \pi]} B_\mu(\omega) \left[2E_\ell(\omega) (2 \nabla E_r \nabla^T E_r + 2 \nabla E_i \nabla^T E_i + \right. \\ & + 2(\mathbf{U}_r(\omega) + \mathbf{U}_r^T(\omega)) \mathbf{h}_\ell \mathbf{U}_r(\omega) \mathbf{h}_\ell + 2(\mathbf{U}_i(\omega) + \mathbf{U}_i^T(\omega)) \mathbf{h}_\ell \mathbf{U}_i(\omega) \mathbf{h}_\ell) + \\ & \left. + 2 \nabla E_\ell(\omega) \nabla^T E_\ell(\omega) \right]. \end{aligned}$$

Для расчёта начальных значений коэффициентов фильтра-прототипа (\mathbf{h}_0) предполагается использовать методы, разработанные для равнополосных банков фильтров [66,67]. Далее на l -й итерации вычисляется вектор невязки при помощи (2.29), после чего обновляется вектор коэффициентов:

$$\mathbf{h}_{\ell+1} = \mathbf{h}_{\ell} + \mathbf{e}_{\ell}.$$

Критерием окончания процедуры является малая величина нормы вектора невязки $\|\mathbf{e}_l\|^2 \leq \eta$, где η — малое положительное число (типичное значение $\eta = 10^{-10}$).

Во внешнем цикле оптимизации происходит перерасчет весовой функции $B_{\mu}(\omega)$. Пусть $h_{\mu,\text{opt}}$ — решение, которое найдено посредством процедуры, описанной выше для конкретного значения функции $B_{\mu}(\omega)$. Вычислим значение функции ошибки $E_{\mu,\text{opt}}(\omega)$ для полученного вектора коэффициентов и определим $V_{\mu}(\omega_l)$ как l -й экстремум функции $|E_{\mu,\text{opt}}(\omega)|$. Далее введем огибающую функцию

$$\beta_{\mu}(\omega) = \frac{\omega - \omega_l}{\omega_{l+1} - \omega_l} V_{\mu}(\omega_{l+1}) + \frac{\omega_{l+1} - \omega}{\omega_{l+1} - \omega_l} V_{\mu}(\omega_l),$$

где l выбирается так, что $\omega_l \leq \omega \leq \omega_{l+1}$. Огибающая функция $\beta_{\mu}(\omega)$ показывает, что больший относительный вес должны иметь те участки в частотной области, на которых $E_{\mu,\text{opt}}(\omega)$ имеет большее значение (рис. 2.13). Новая весовая функция $\beta_{\mu+1}(\omega)$ получается с использованием следующего правила [66]:

$$B_{\mu+1}(\omega) = B_{\mu}(\omega) \frac{\beta_{\mu}^{\theta}(\omega)}{A_{\mu}},$$

где нормирующий множитель A_{μ} определяется выражением

$$A_{\mu} = \sum_{\omega \in [0, \pi]} B_{\mu}(\omega)^2 \sqrt{\sum_{\omega \in [0, \pi]} \beta_{\mu}(\omega)^{2\theta}}.$$

Здесь θ — положительное число, влияющее на сходимость алгоритма. Начальное значение весовой функции $B_0(\omega) = 1, \forall \omega \in [0, \pi]$.

Итерационный алгоритм завершается, когда

$$\frac{\max \beta_{\mu}(\omega) - \min \beta_{\mu}(\omega)}{\max \beta_{\mu}(\omega) + \min \beta_{\mu}(\omega)} \leq \psi,$$

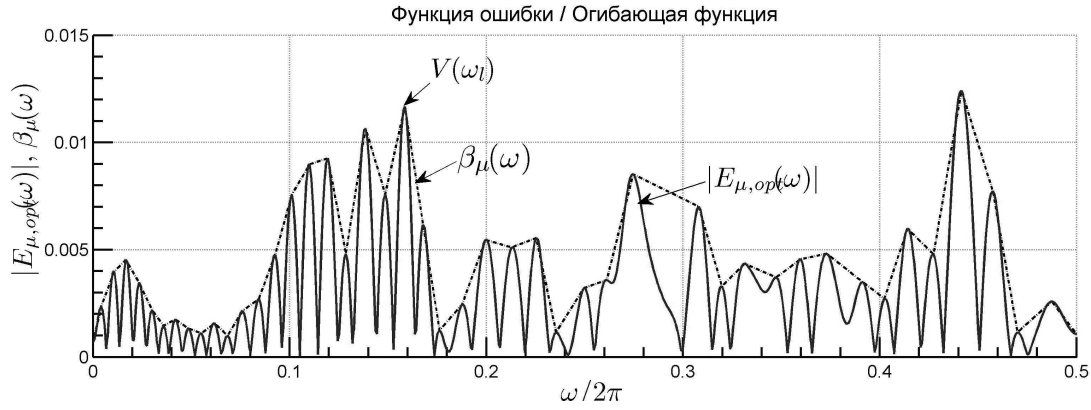


Рис. 2.13. Функция ошибки $|E_{\mu,\text{opt}}(\omega)|$ и её огибающая $\beta_{\mu}(\omega)$

где ψ – константа, определяющая «пологость» огибающей $\beta_{\mu}(\omega)$. Ориентировочные значения для параметров внешнего цикла оптимизации $0,5 \leq \psi \leq 0,6$ и $1 \leq \theta \leq 1,5$. Число точек частотной сетки для расчета функции ошибки, как правило, выбирается равным $4N$, где N – порядок фильтра-прототипа.

2.4. Результаты экспериментов

2.4.1. Расчет коэффициентов децимации/интерполяции

В качестве практического примера рассмотрим 18-канальный неравнополосный КМБФ, который аппроксимирует психоакустическую шкалу барков при частоте дискретизации сигнала $f_s = 8$ кГц. Для этого, используя (2.21), установим коэффициент деформации оси частот $\alpha = 0,3751$. Далее воспользуемся алгоритмом определения частотных полос неравнополосного КМБФ для нахождения коэффициентов децимации/интерполяции. В качестве примера определим коэффициент децимации/интерполяции для третьего канала ($k = 2$) банка фильтров. Необходимые границы частот равнополосного банка фильтров равны

$$F_1^u = \frac{\pi}{18}; \quad F_4^u = \frac{4\pi}{18}.$$

Используя (2.23), найдем соответствующий частотный диапазон для третьего канала НКМБФ (значения частот нормированы к значению частоты

дискретизации):

$$f_{L_2} = \frac{\varphi(F_1^u)}{2\pi} \approx 0,0119; \quad f_{U_2} = \frac{\varphi(F_4^u)}{2\pi} \approx 0,0493.$$

Применим правило (2.22) для определения коэффициента S_2 :

$$\begin{aligned} 1 \leq n_2 \leq \left\lfloor \frac{f_{U_2}}{f_{U_2} - f_{L_2}} \right\rfloor &= \lfloor 1,3196 \rfloor = 1 \Rightarrow n_2 = 1; \\ \left\lfloor \frac{n_2}{2f_{U_2}} \right\rfloor \geq S_2 \geq \left\lceil \frac{n_2 - 1}{2f_{L_2}} \right\rceil &\Rightarrow \left\lfloor \frac{1}{0,0985} \right\rfloor \geq S_2 \geq \left\lceil \frac{0}{0,0239} \right\rceil \Rightarrow \\ &\Rightarrow [10,15] \geq S_2 \geq [0] \Rightarrow S_2 = 10. \end{aligned}$$

Аналогичным образом определяются оставшиеся коэффициенты децимации/интерполяции канальных сигналов неравнополосного КМБФ:

$$S_k = \{28, 13, 10, 7, 6, 5, 9, 7, 6, 5, 4, 4, 3, 3, 1, 2, 2, 4\}, \quad 0 \leq k < 18.$$

По этим значениям можно вычислить коэффициент передискретизации банка фильтров [78] (*oversampling factor*), иногда называемый также общим отношением передискретизации (*total oversampling ratio*), который рассчитывается как

$$\mathcal{O} = \sum_{k=0}^{M-1} \frac{1}{S_k}. \quad (2.30)$$

Данный показатель отражает отношение числа отсчетов сигналов в каналах банка фильтров к числу отсчетов входного сигнала. В случае критически децимированного банка фильтров это отношение равно единице, что означает отсутствие информационной избыточности (последовательность из M входных отсчетов преобразуется в M канальных сигналов – по одному отсчету в каждом канале). Для рассматриваемого банка фильтров $\mathcal{O} = 4,75$, что говорит о наличии определенной информационной избыточности сигналов в каналах банка фильтров. Уменьшение значения \mathcal{O} привело бы к увеличению уровня искажений, образующихся в результате эффекта наложения спектров. Значение $\mathcal{O} > 1$ обычно является платой за уменьшение уровня искажений, вносимых банком фильтров. Так, для банка фильтров, предлагаемого в [78], общее отношение передискретизации равно четырем.

2.4.2. Оптимизация фильтра-прототипа

В качестве начального приближения рассчитаем фильтр-прототип для банка фильтров, используя метод [66], порядок фильтра-прототипа примем $N = 144$. В результате предварительного расчета получен неравнополосный банк фильтров анализа (рис. 2.14).

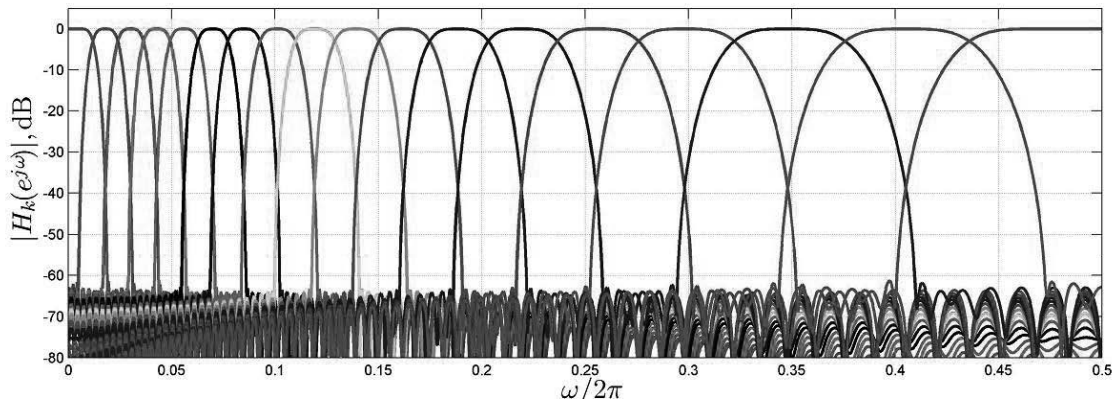


Рис. 2.14. АЧХ неравнополосного банка фильтров анализа ($M = 18$, $\alpha = 0,3751$)

Результат применения предлагаемого метода проектирования неравнополосного КМБФ показан на рис. 2.15. По виду частотной характеристики оптимизированного фильтра-прототипа можно сделать вывод, что оптимизационная процедура увеличила коэффициент ослабления в тех частотных областях, которые влияют на подавление компонент наложения спектров, появляющихся как результат процесса децимации/интерполяции сигналов в каналах банка фильтров.

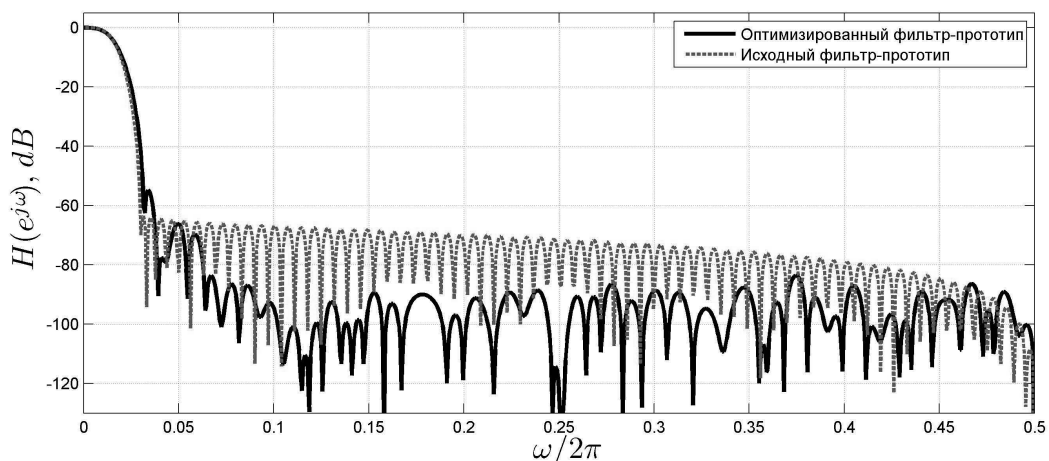


Рис. 2.15. АЧХ фильтров-прототипов неравнополосного КМБФ

Результаты сравнения функции $T_{\text{all}}(e^{j\omega})$ исходного и оптимизированного банков фильтров (рис. 2.16) показывают, что предлагаемый метод проектирования позволяет уменьшить уровень искажений с 0,018 до 0,002 дБ.

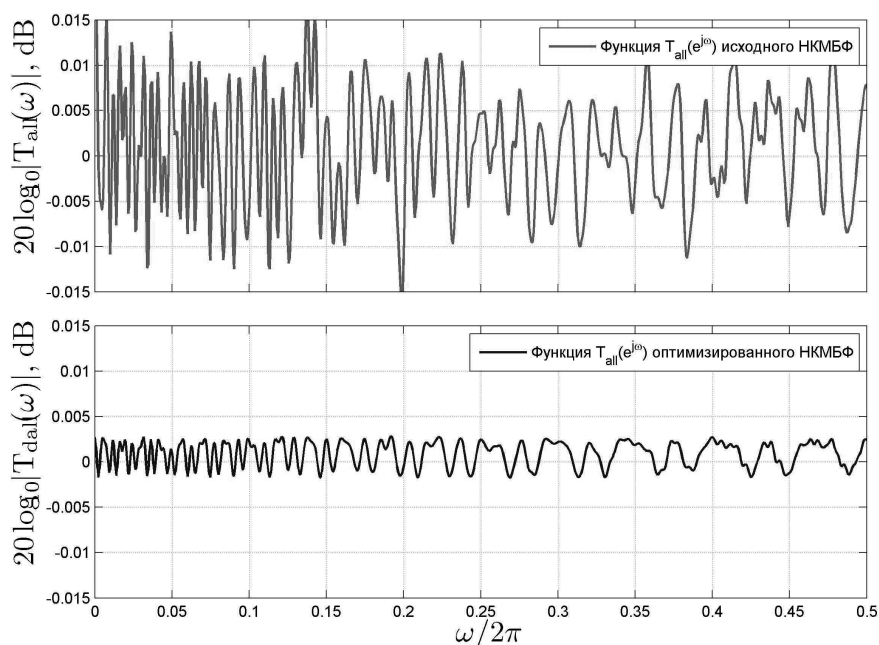


Рис. 2.16. Частотные характеристики исходного и оптимизированного НКМБФ

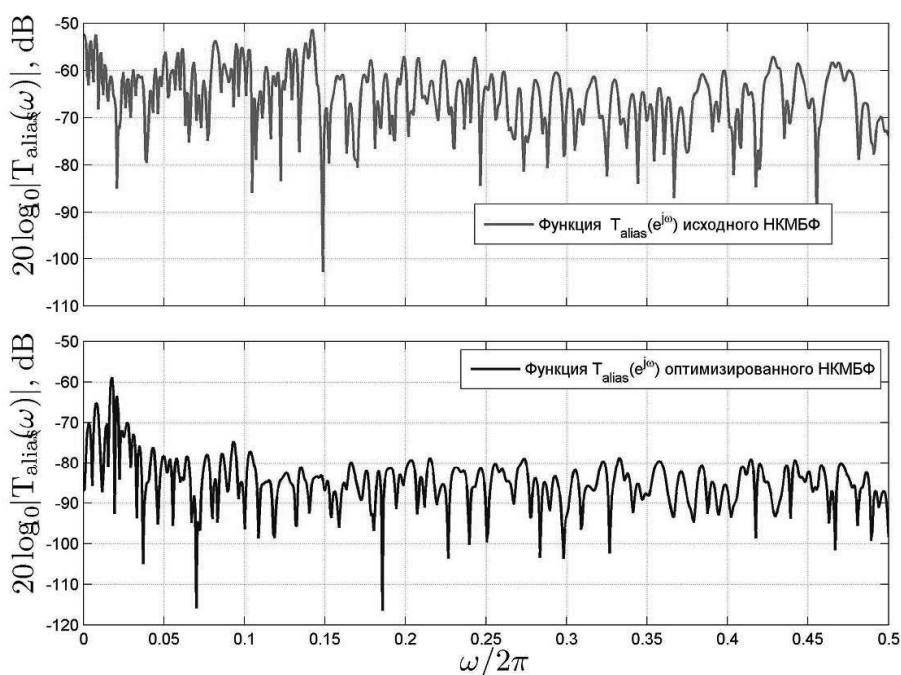


Рис. 2.17. Частотная характеристика передачи компонент наложения спектров исходного и оптимизированного НКМБФ

Рисунок 2.17 дает представление об уровне компонент наложения спектров, присутствующих в реконструированном сигнале. В среднем уровень компонент наложения спектра уменьшается на 20 дБ.

2.4.3. Бичастотная функция неравнополосного КМБФ

Лучшее представление о возникающих в системе анализа/синтеза банка фильтров нежелательных компонентах наложения спектров дает построение *бичастотной функции системы* с переменными параметрами (*time-varying system*). Линейная система с переменными параметрами может быть записана в виде следующего соотношения вход/выход [79]:

$$y[m] = \sum_{n=-\infty}^{\infty} K[m, n]x[n],$$

где $K(m, n)$ — отклик системы (или функция Грина), которая соответствует отклику системы в момент времени m на входное воздействие в момент времени n . Функцию

$$K(e^{j\omega'}, e^{j\omega}) = \frac{1}{2\pi} \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} K[m, n]e^{j\omega'n}e^{j\omega m}$$

называют бичастотной функцией системы. Данная функция определяет отображение из частотной области сигнала $X(e^{j\omega})$ в частотную область $Y(e^{j\omega})$.

Бичастотная функция позволяет оценить величину компонент наложения АЧХ каналов банка фильтров, возникающих в результате процесса децимации/интерполяции сигнала, в каждом отдельном канале. На рис. 2.18–2.19 изображены компоненты наложения спектров, появляющиеся в результате процесса децимации/интерполяции только во втором канале банка фильтров ($S_1 = 13$). В следующем примере (рис. 2.20–2.21) показываються бичастотные функции для исходного и оптимизированного неравнополосного КМБФ (при децимации/интерполяции всех каналов). Через ω обозначена круговая частота входного сигнала, а через ω' — круговая частота реконструированного сигнала. Бичастотная функция показывает, как осуществляется перенос энергии в частотной области между входным

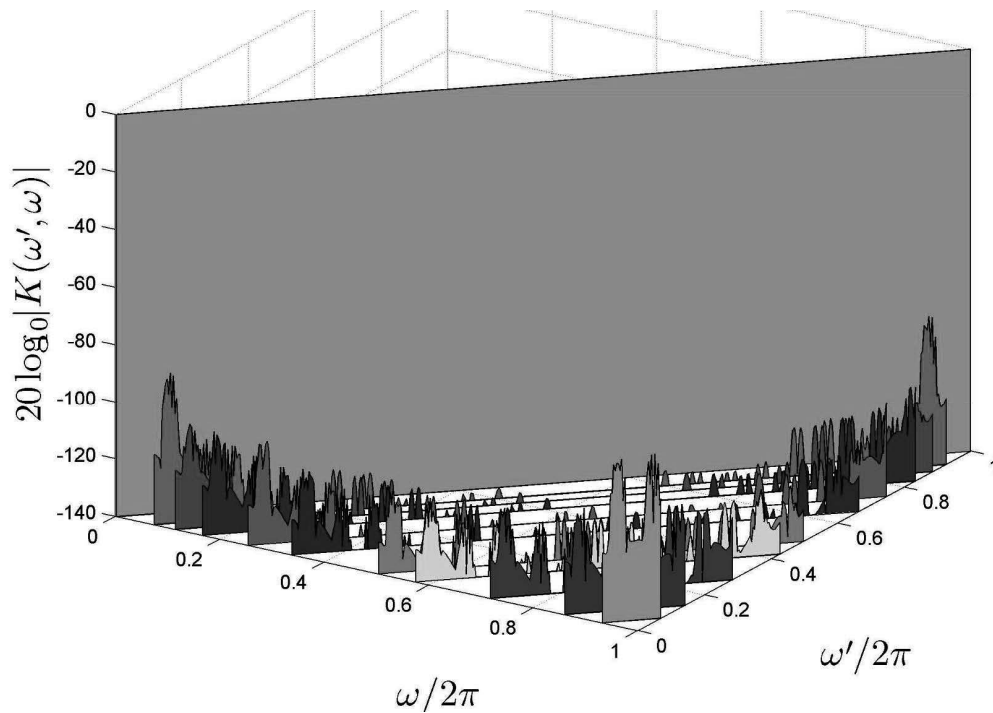


Рис. 2.18. Компоненты наложения спектров, возникающие в результате децимации/интерполяции во втором канале НКМБФ: исходный банк фильтров

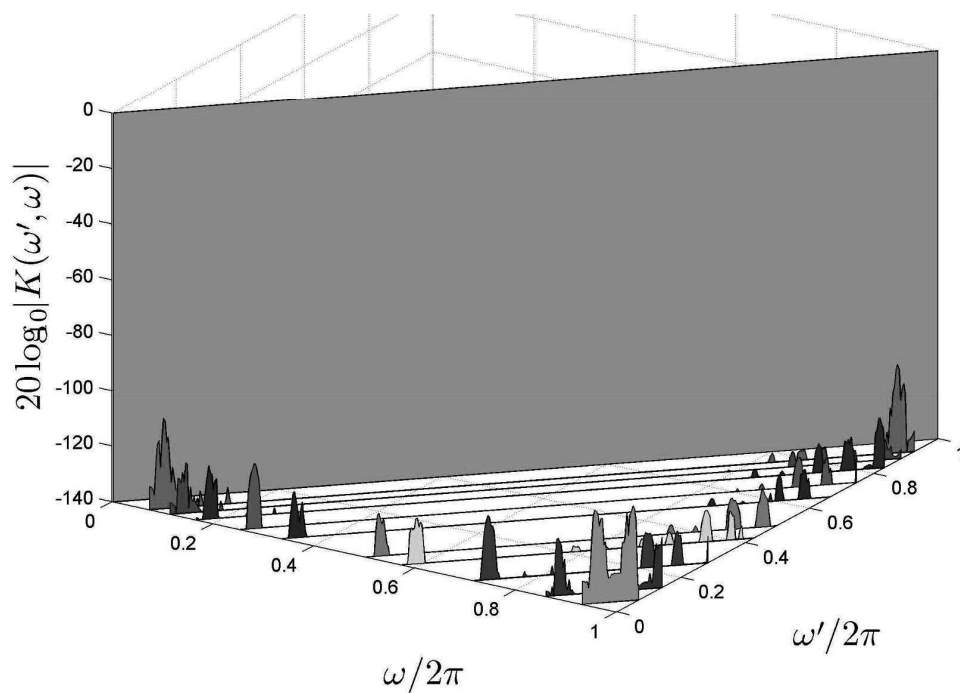


Рис. 2.19. Компоненты наложения спектров, возникающие в результате децимации/интерполяции во втором канале НКМБФ: оптимизированный банк фильтров

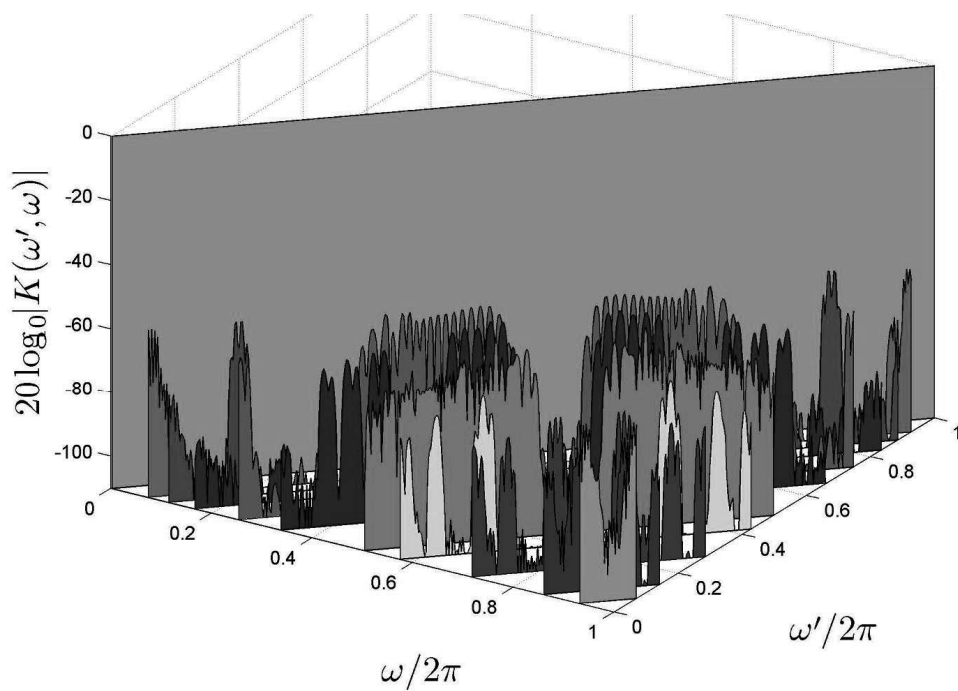


Рис. 2.20. Бичастотная функция исходного неравнополосного КМБФ

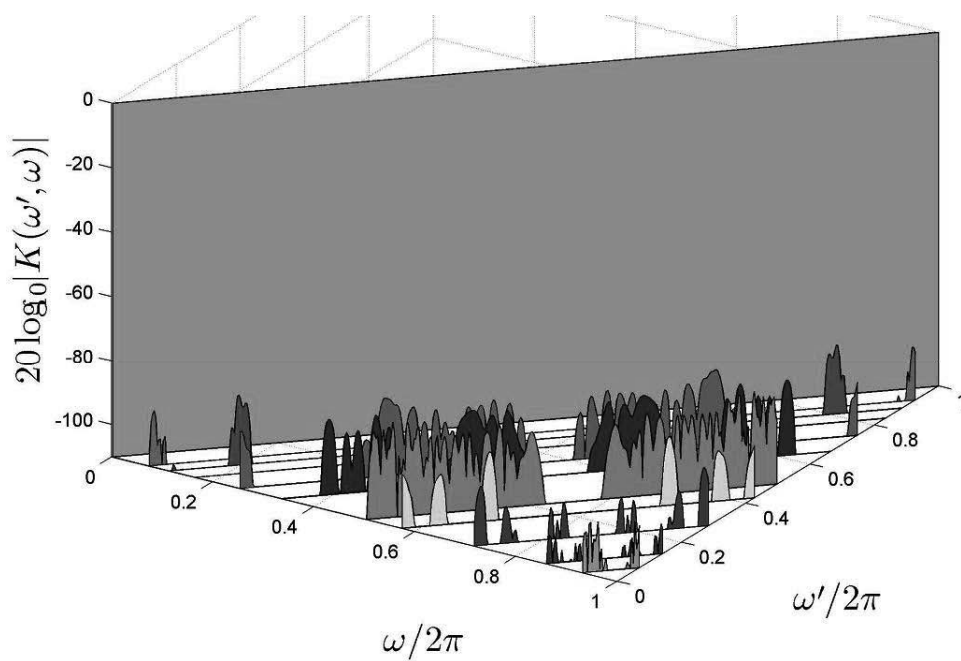


Рис. 2.21. Бичастотная функция оптимизированного неравнополосного КМБФ

и реконструированным сигналами. Как видно из рисунка, основная энергия сигнала остается на исходных частотах (главная диагональ).

Бичастотная функция оптимизированного неравнополосного КМБФ (рис. 2.21) показывает, что уровень компонент наложения спектров, появляющихся в результате процесса децимации/интерполяции, значительно ниже, чем у исходного неравнополосного КМБФ. Это объясняется тем, что на этапе оптимизации фильтра-прототипа учитывался специфический набор коэффициентов децимации/интерполяции S_k . Полученный в результате оптимизации фильтр-прототип (и банк фильтров на его основе) позволяет лучше ослаблять компоненты наложения спектров, что приводит к меньшему уровню искажений в реконструированном сигнале.

ГЛАВА 3

ПЕРЕКРЫВАЮЩИЕСЯ ПРЕОБРАЗОВАНИЯ И ИХ ПРИМЕНЕНИЕ В СЛУХОВЫХ АППАРАТАХ

3.1. Причины применения перекрывающихся преобразований в слуховых аппаратах

Одним из основных требований, предъявляемым к схеме обработки сигнала в слуховом аппарате, является малая групповая задержка. В первую очередь это связано с эффектом эха, который возникает в результате суперпозиции обработанного (задержанного) сигнала с необработанным сигналом, который попадает в ушной канал напрямую через отверстие (вент) в слуховом аппарате. Данный эффект можно устранить, если использовать вкладыш, плотно закрывающий ушной проход. Однако это приводит к другому нежелательному последствию – эффекту *окклюзии*. В результате закупоривания внешней части ушного канала низкочастотные вибрации, появляющиеся в момент разговора не могут покинуть полости уха и от этого усиливаются. На низких частотах усиление может достигать 20 дБ, а собственный голос пациента будет казаться ему звучащим неестественно. Избежать эффекта окклюзии помогает вентиляционное отверстие в ушном вкладыше, которое, однако, приводит к риску появления эффекта акустической обратной связи. Другим аспектом работы слухового аппарата является эффект эха, который возникает в момент произнесения слов пользователем слухового аппарата. Собственный голос достигает барабанной перепонки, как посредством костной проводимости, так и через слуховой аппарат. Таким образом, эффекта эха можно избежать, только уменьшив групповую задержку алгоритма обработки сигнала в слуховом аппарате. По оценкам современных исследователей общая задержка сигнала в слуховом аппарате не должна превышать 6–8 мс [5]. Это время включает в себя задержку на преобразование сигнала из аналоговой формы в цифровую и обратно, которая в настоящее время оценивается равной 0,5–2 мс [6]. Одним из возможных путей уменьшения групповой задержки

сигнала в слуховом аппарате является использование перекрывающегося преобразования для частотной декомпозиции.

3.2. Перекрывающееся преобразование (ПП)

Перекрывающееся преобразование (*lapped transform*), особенно модифицированное ДКП (МДКП), широко применяется в высококачественном сжатии аудиосигналов [54]. Однако уже в [80] было показано, что перекрывающееся преобразование (ПП) может эффективно применяться в слуховых аппаратах для организации обработки сигнала в частотной области. Преимущество ПП в том, что оно позволяет уменьшить (либо полностью исключить) блочный эффект (*blocking effect*), образующийся вследствие отсутствия непрерывности на границах последовательных блоков [81], который характерен для большинства блочных преобразований. При кодировании речи блочный эффект приводит к появлению посторонних тонов [82]. В тоже время ПП имеет эффективную с вычислительной точки зрения реализацию, что также очень важно для применения в слуховых аппаратах.

Общепринятым подходом к устранению блочного эффекта является обработка сигнала перекрывающимися сегментами. Из-за того, что сегменты имеют перекрывающиеся участки, общее количество отсчетов, которое необходимо обрабатывать, возрастает. В случае кодирования сигнала это обстоятельство приводит к увеличению скорости потока данных. Преимущество ПП в том, что оно позволяет устранить блочный эффект и не приводит к увеличению количества данных (для обработки или кодирования). В общем случае перекрывающееся ортогональное преобразование характеризуется тем, что каждый блок длиной M отсчетов отображается в набор из M базисных функций, длина каждой из которых больше M .

3.3. Дополненное модулированное ПП (ДМПП)

В [83] предложено использовать дополненное модулированное перекрывающееся преобразование (ДМПП) (*AML*P – *augmented modulated*

lapped transform) для построения субполосной системы обработки сигнала слухового аппарата.

3.3.1. Модулированное перекрывающееся преобразование

Базисные функции модулированного перекрывающегося преобразования (МПП) имеют длину $L = 2M$ и записываются следующим образом:

$$p_k[n] = h[n] \cos \left[\left(k + \frac{1}{2} \right) \left(n + \frac{M+1}{2} \right) \frac{\pi}{M} \right],$$

где $n = 0, 1, \dots, L-1$ – временной индекс;

$k = 0, \dots, M-1$ – индекс субполосы.

Фильтр-прототип $h[n]$ определяется как

$$h[n] = -\sin \left[\left(n + \frac{1}{2} \right) \frac{\pi}{L} \right]. \quad (3.1)$$

Это единственный допустимый фильтр нижних частот, который удовлетворяет требованиям полной реконструкции и полифазной нормализации [81]. Базисные функции МПП имеют ослабление первого бокового лепестка в 23 дБ, в то время как ослабление первого бокового лепестка базисных функций ДПФ равно 13 дБ. Ослабление боковых лепестков МПП возрастает со скоростью 40 дБ/декаду и имеет 60 дБ ослабления для нормализованной частоты Ω , отстоящей от центральной частоты Ω_C более чем на $\pi/4$ (рис. 3.1).

Значение сигнала в k -й субполосе в m -м блоке МПП определяется как

$$X_k[m] = \sqrt{\frac{2}{M}} \sum_{r=0}^{L-1} x_m[r] p_k[r],$$

где $x_m[r] = x[mM - L + 1 + r]$.

3.3.2. Определение ДМПП

Дополненное МПП образуется путем присоединения к базисной функции $p_k[n]$ комплексной части для того, чтобы иметь возможность впоследствии вычислять фазовый сдвиг. Таким образом, базисная функция

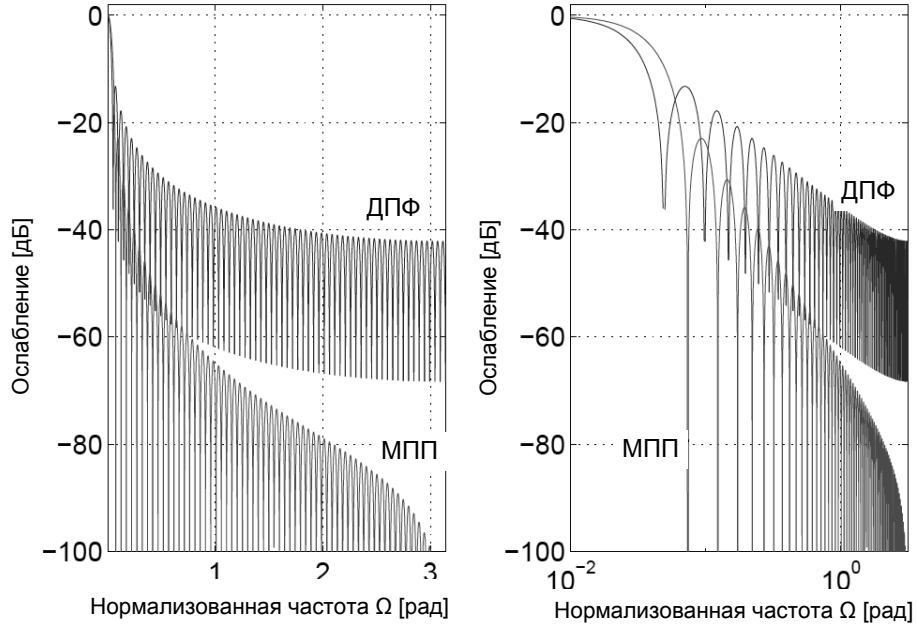


Рис. 3.1. Сравнение базисных функций ДПФ и МПП ($k = 0$, $M = 64$)

$p_k^C[n]$ ДМПП определяется как

$$p_k^C[n] = p_k[n] + j \cdot h[n] \sin \left[\left(k + \frac{1}{2} \right) \left(n + \frac{M+1}{2} \right) \frac{\pi}{M} \right].$$

Используя данное определение ДМПП¹, k -й субполосный сигнал $X_k^{C,0}[m]$ вычисляется как

$$X_k^{C,0}[m] = \sqrt{\frac{2}{M}} \sum_{r=0}^{L-1} x_m[r] h[r] e^{j \frac{\pi}{M} (k + \frac{1}{2})(r + \frac{M+1}{2})}, \quad (3.2)$$

где $\text{Re}\{X_k^{C,0}[m]\} = X_k[m]$.

Выражение (3.2) можно переписать в следующем виде:

$$X_k^{C,0}[m] = \sqrt{\frac{2}{M}} \underbrace{\left(\sum_{r=0}^{L-1} u_m[r] W_L^{r(k + \frac{1}{2})} \right)}_{X_k^{C,1}[m]} \underbrace{W_L^{(k + \frac{1}{2})(\frac{M+1}{2})}}_{\phi(k)}$$

где $W_L(r) = e^{j \frac{2\pi}{L} r}$; $u_m[r] = x_m[r] h[r]$. Интерпретация $X_k^{C,1}[m]$ позволяет получить следующую полифазную структуру ДМПП (рис. 3.2).

В структуре на рис. 3.2 используется симметричное ДПФ, которое

¹с учетом формулы Эйлера $e^{j\varphi} = \cos \varphi + j \sin \varphi$

определяется как

$$\text{Симметричное ДПФ : } F[k] = \sum_{n=0}^{L-1} f[n] W_L^{n(k+\frac{1}{2})}, \quad k = 0, \dots, M-1.$$

Также приведенная схема выявляет тесную связь между ДМПП и ДПФ-модулированным банком фильтров.

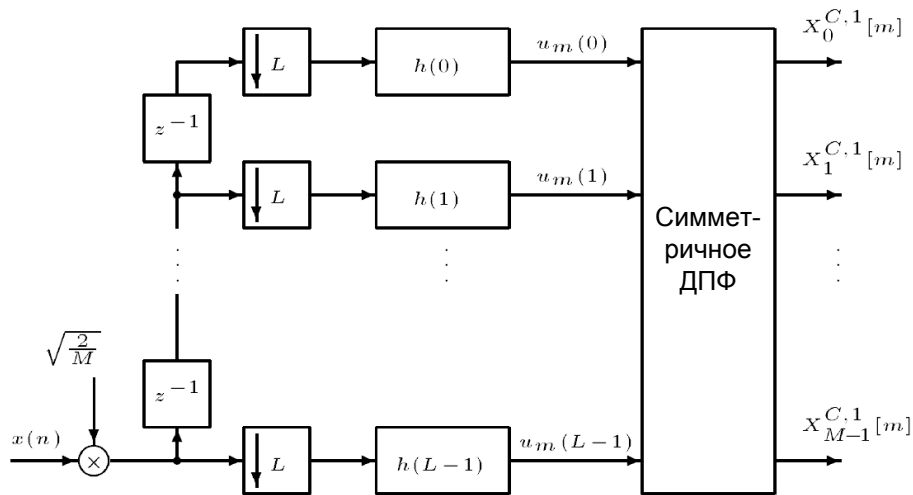


Рис. 3.2. Банк фильтров анализа (без поворачивающих множителей $\phi(k)$)

3.3.3. Схема компенсации акустической обратной связи на основе ДМПП

На рис. 3.3 приведена схема слухового аппарата, задачей которого является усиление акустического сигнала $v(t)$ для компенсации потери слуха тугоухого человека. Вследствие акустической и механической обратной связи микрофон улавливает не только входной сигнал $v(t)$, но и сигнал обратной связи $y(t)$. В большинстве практических реализаций [84] алгоритм коррекции (корректор) вносит временную задержку, обозначенную на схеме через T .

Система содержит устройство клипирования, которое не позволяет сигналу превысить уровень насыщения динамика. За счет введения этого устройства *линейный* предсказатель избавляется от необходимости оценивать *нелинейность* канала обратной связи [85]. Для моделирования канала акустической обратной связи, который включает в себя АЦП, ЦАП, воздушный канал, микрофон и динамик, можно использовать КИХ-фильтр

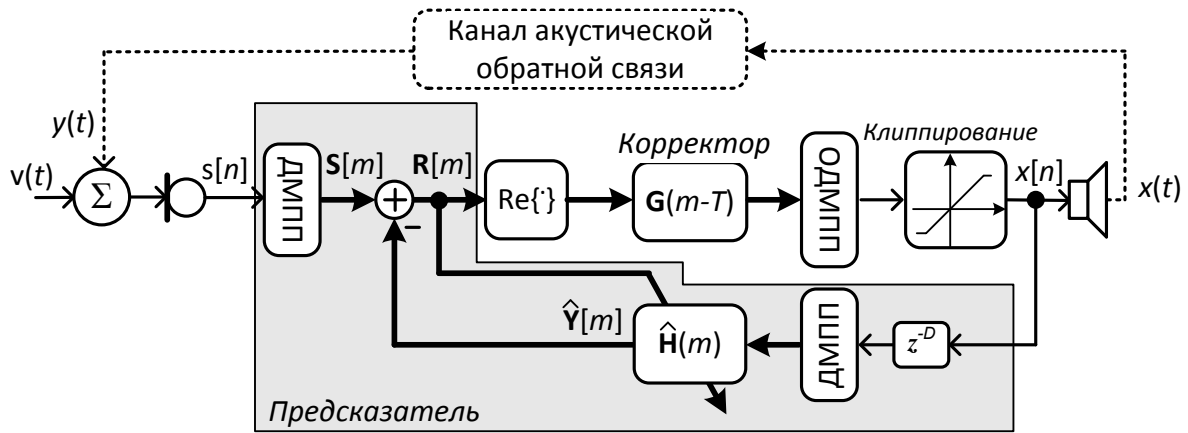


Рис. 3.3. Схема компенсации акустической обратной связи на основе ДМП (n – номер отсчета, m – номер блока)

$h_{afp}[n]$ порядка N . Задержка на D отсчетов в схеме на рис. 3.3 вводится для сдвига «центра тяжести» импульсной характеристики предсказателя. В качестве примера на рис. 3.4 приведены импульсные характеристики каналов обратной связи заушного (*BTE* – *behind-the-ear*) и внутриканального (*ITE* – *in-the-ear*) слуховых аппаратов. Все импульсные характеристики нормированы так, что $\sum_n h_{afp}[n]^2 = 1$. В [86] предлагается определять эффективную длину импульсной характеристики канала акустической обратной связи N_e как

$$N_e = \arg \min_N \left\{ \sum_{r=0}^N h_{afp}[r]^2 > 0,9 \right\}.$$

Для импульсных характеристик на рис. 3.4 эффективная длина N_e имеет следующие значения: 25, 27, 13 и 25. Можно отметить, что длина канала акустического обратной связи слуховых аппаратов ВТЕ типа в два раза больше, чем у слухового аппарата ITE типа, за исключением «патологического» случая (ITE №2), где длина канала в слуховом аппарате ITE типа такая же, как и в слуховом аппарате ВТЕ типа.

3.3.4. Предсказатель $\hat{H}[m]$

Сигнал компенсации эхо-сигнала $\hat{Y}[m]$ вычисляется как

$$\hat{Y}[m] = \mathbf{X}[m]\hat{H}[m],$$

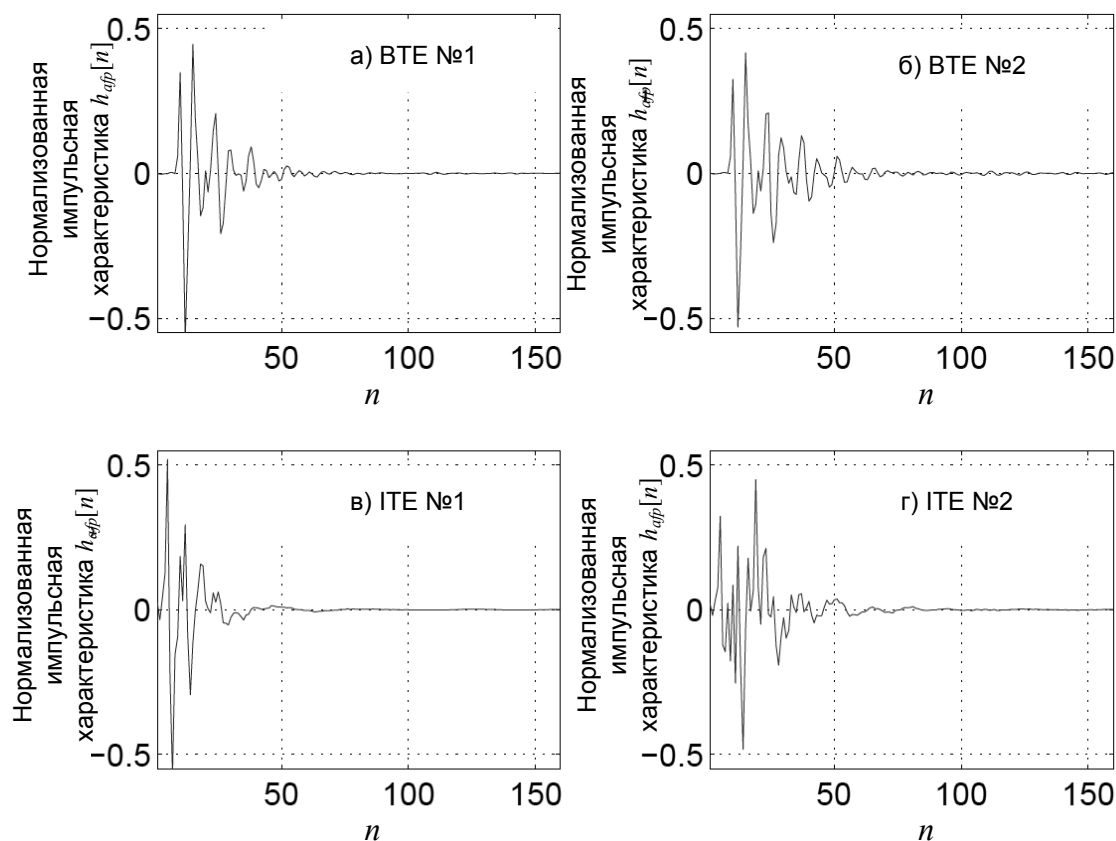


Рис. 3.4. Импульсная характеристика канала акустической обратной связи слухового аппарата (частота дискретизации 16 кГц): а) ВТЕ; б) ВТЕ, находящийся на расстоянии 10 см от деревянной перегородки, параллельной уху; в) ITE; г) ITE с поднесенной рукой, закрывающей ухо

где $\mathbf{X}[m] = \text{diag}_{0 \leq k < M}(X_k^{C,0}[m])$ – входная матрица для предсказателя. Коэффициенты фильтра-предсказателя обновляются согласно алгоритму NLMS (NLMS – normalized least mean square):

$$\hat{H}[m+1] = \hat{H}[m] + \mu[m]R[m],$$

где $\mu[m]$ – диагональная матрица, отвечающая за скорость обучения;

$$R[m] = S[m] - \hat{Y}[m] - \text{сигнал ошибки.}$$

Элементы матрицы $\mu[m]$ определяются следующим образом:

$$\mu_{k,k}[m] = \mu_0 / E\{|X_k^{C,0}|^2\}.$$

При практическом использовании оператор математического ожидания $E\{\cdot\}$ реализуют путем экспоненциального усреднения (с коэффициентом забывания γ).

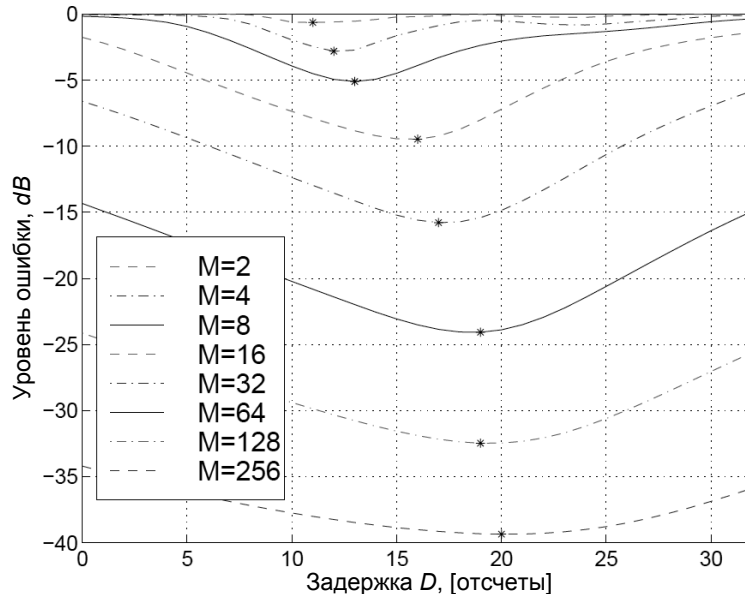


Рис. 3.5. Уровень ошибки как функция задержки D и количества каналов M

В [86] оптимальное значение задержки D (в смысле наилучшей компенсации обратной связи) предложено искать путем преобразования сигнала $y[n] = h_{afp}[n] * x[n]$ и $x[n - D]$ в область коэффициентов ДМПП и решением уравнения (в среднеквадратичном смысле)

$$Y[m] = \mathbf{X}[m]\hat{H}$$

относительно \hat{H} для нескольких блоков m . Уровень ошибки определяется как $\frac{\|r[n]\|}{\|x[n]\|}$, где сигнал ошибки $r[n]$ определяется путем обратного ДМПП от $R[m]$. На рис. 3.5 показано, что с увеличением числа полос M уровень ошибки уменьшается, а кривая зависимости становится более полой и, следовательно, влияние задержки D на производительность предсказателя значительно уменьшается при $N_e \ll M$.

3.4. Модифицированное дискретное косинусное преобразование с деформацией частотной оси

В работе [87] исследовалась возможность применения модифицированного косинусного преобразования (МДКП) с деформированной частотной осью для обработки сигнала в слуховом аппарате. Особенности преобразования являются разложение сигнала на частотные составляющие с

неравной шириной полос в соответствии с особенностями слуховой системы человека, а также малая групповая задержка.

Модифицированное ДКП (МДКП), которое является разновидностью перекрывающегося преобразования, особенно широко применяется в высококачественном сжатии аудиосигналов [54]. В этом приложении требуются хорошая частотная избирательность и отсутствие избыточности в субканалах, поэтому для кодирования используют длинные временные окна (до 2048 коэффициентов) и критическую децимацию субканальных сигналов. Предполагается, что для слуховых аппаратов можно получить более простое перекрывающееся преобразование с небольшим числом субканалов, в которых отсутствует децимация. Это позволит избежать проблем с наложением спектра вследствие передискретизации. Главным недостатком подхода является увеличение вычислительной сложности, которая, однако, может быть уменьшена на уровне реализации. Алгоритм субполосной декомпозиции в слуховых аппаратах не обязан обладать свойством полного восстановления, что позволяет использовать фильтр-прототип с большим ослаблением в полосе заграждения.

3.4.1. Модифицированное дискретное косинусное преобразование (МДКП)

МДКП определяется следующим образом:

$$y[k] = \sum_{n=0}^{2M-1} h[n]x[n] \cos\left(\frac{\pi(2n + (M + 1))(2k + 1)}{4M}\right), \quad (3.3)$$

где $k = 0, \dots, M - 1$ – номер частотного канала;

$h[n]$ – фильтр-прототип.

Часто в качестве фильтра-прототипа выбирают следующую функцию:

$$h[n] = \sin\left(\frac{\pi(n + 1/2)}{2M}\right), \quad n = 0, \dots, 2M - 1. \quad (3.4)$$

Известно, что МДКП можно эффективно реализовать при помощи дискретного косинусного преобразования четвертого типа (ДКП-4). Для

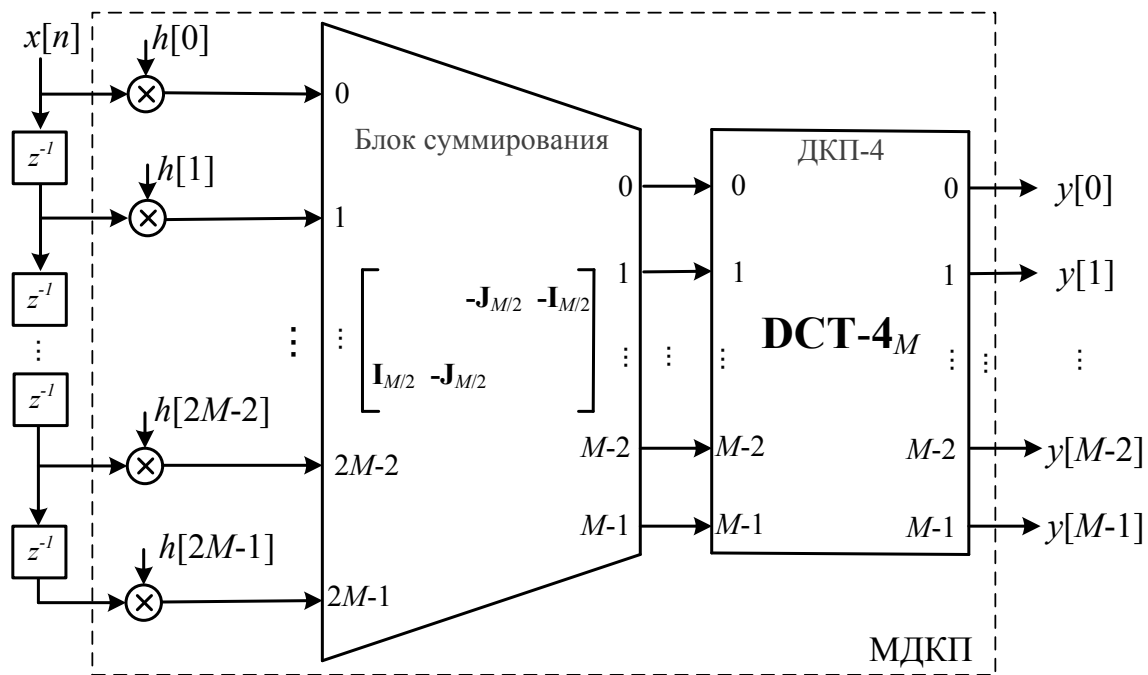


Рис. 3.6. Граф-схема алгоритма МДКП

этого (3.3) необходимо представить в векторно-матричной форме:

$$\mathbf{y} = \text{DCT-4}_M \begin{bmatrix} & -J_{\frac{M}{2}} & -I_{\frac{M}{2}} \\ I_{\frac{M}{2}} & -J_{\frac{M}{2}} & \end{bmatrix} \cdot \text{diag}(h[0], \dots, h[2M-1]) \cdot \mathbf{x}, \quad (3.5)$$

где DCT-4_M – матрица ДКП-4 размера $M \times M$;

$\mathbf{x} = [x[0], \dots, x[2M-1]]$ – вектор входных данных;

$\mathbf{y} = [y[0], \dots, y[M-1]]$ – вектор выходных данных.

На основании выражения (3.5) строится следующая граф-схема алгоритма МДКП (рис. 3.6). Каждый новый отсчет поступает в линию задержки, в которой хранятся последние $2M$ входные отсчеты, затем выполняется умножение на коэффициенты фильтра-прототипа. Результат умножения обрабатывается в блоке суммирования, выход которого поступает в блок ДКП-4. Таким образом, МДКП можно рассматривать как частный случай равнополосного косинусно-модулированного банка фильтров [75]. Приведенная граф-схема является эффективной, поскольку разработаны быстрые алгоритмы вычисления ДКП-4 [88–91], имеющие вычислительную сложность $O(\frac{M}{2} \log_2 M)$.

3.4.2. Искривление частотной оси МДКП

Разрешающая способность слуховой системы человека нелинейна и ухудшается с повышением частоты [11]. Поэтому в слуховых аппаратах требуется неравнополосный банк фильтров, отвечающий данному свойству. Чтобы получить из МДКП неравнополосный банк фильтров, предлагается использовать фазовое преобразование [74], которое заключается в замене каждого элемента задержки z^{-1} на фазовое звено $A(z)$ (см. подраздел 2.2.2.). В [48] показано, что для аппроксимации психоакустических шкал достаточно использования фазового звена первого порядка.

На рис. 3.7 приведены частотные характеристики обычного МДКП и его неравнополосной версии. Частота дискретизации предполагалась равной 16 кГц. Параметр деформации частотной оси α выбирался таким образом, чтобы МДКП аппроксимировало психоакустическую шкалу барков. Число каналов $M=22$.

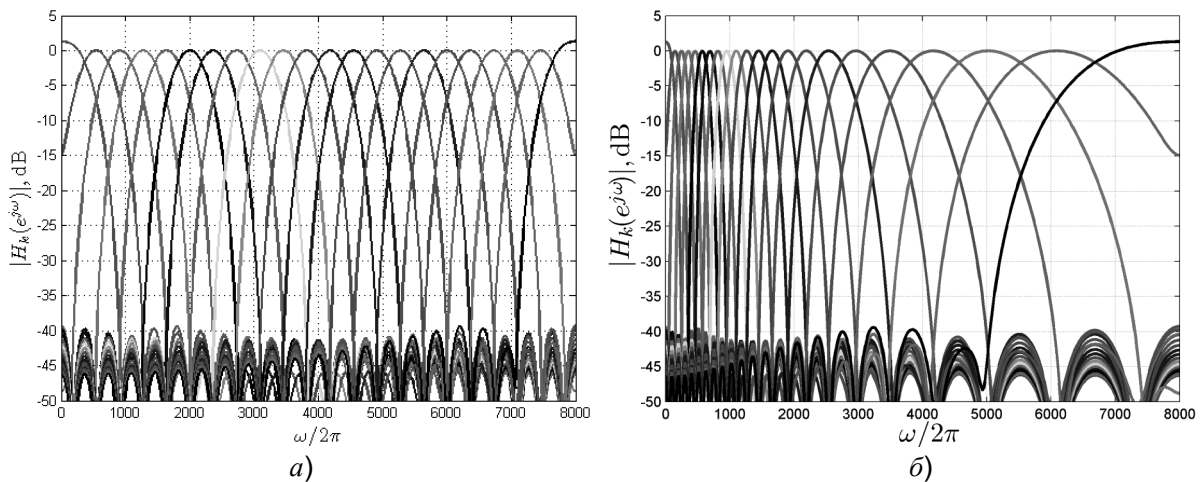


Рис. 3.7. Частотная характеристика МДКП: а) обыкновенного, б) с деформацией частотной оси

3.4.3. Применение МДКП с деформацией частотной оси в слуховом аппарате

На рис. 3.8 показано применение МДКП с деформированной частотной осью в слуховом аппарате. Коэффициенты g_0, \dots, g_{M-1} используются для формирования заданного частотного профиля. Поскольку сигналы в каналах не децимируются, то выполнить синтез выходного сигнала можно

путем суммирования всех субканальных сигналов. В процессе работы слухового аппарата, коэффициенты усиления g_0, \dots, g_{M-1} могут изменяться в зависимости от энергии сигнала в каждой субполосе.

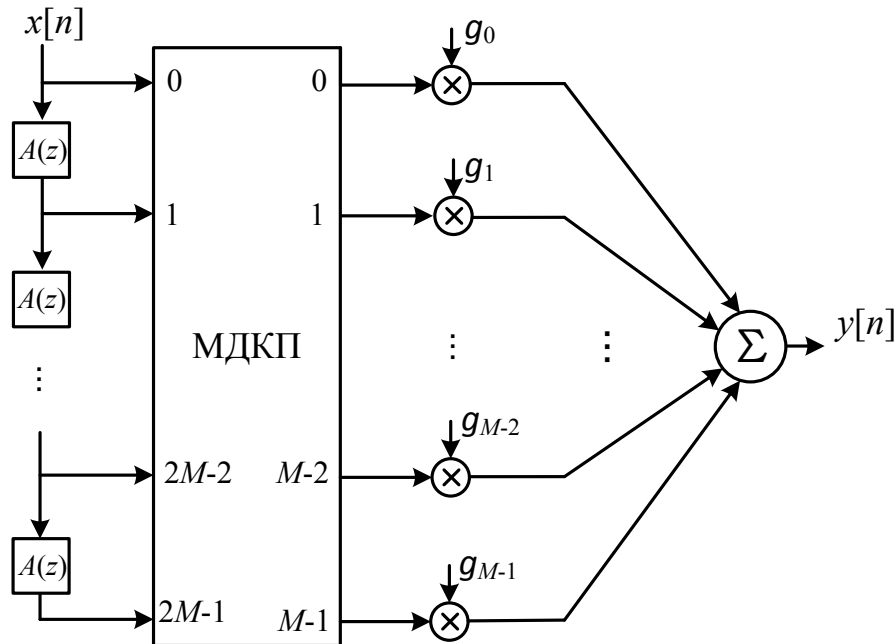


Рис. 3.8. Применение МДКП с деформированной частотной шкалой в слуховом аппарате

Особенность предлагаемой схемы (рис. 3.8) в том, что она значительно сокращает групповую задержку сигнала, что, как отмечалось, весьма важно для слуховых аппаратов. Так, для частоты дискретизации 16 кГц групповая задержка системы на рис. 3.8 составляет порядка 3 мс.

В слуховом аппарате при помощи коэффициентов g_0, \dots, g_{M-1} , как правило, необходимо достичь заданной формы спектра выходного сигнала. Примером может служить компенсация потери слуха, которая производится при помощи тональных аудиограмм [13]. По имеющейся аудиограмме для каждой частоты рассчитываются коэффициенты компенсации потери слуха (*insertion gain*). Спектральная форма функции компенсации потери слуха зависит от метода корректировки слуха и не обязательно полностью повторяет форму аудиограммы [6]. На рис. 3.9 показаны примеры аудиограмм и рассчитанные для них частотные профили компенсации потери слуха. Также на рисунке показано, как предлагаемое МДКП с деформированной частотной осью аппроксимирует заданный профиль компенсации потери слуха.

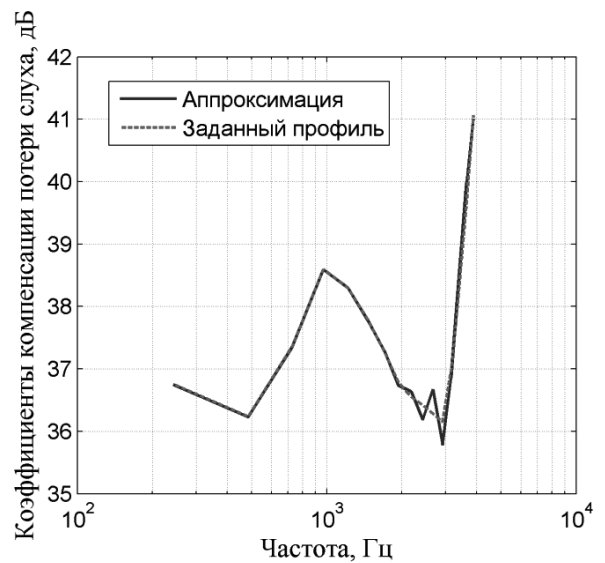
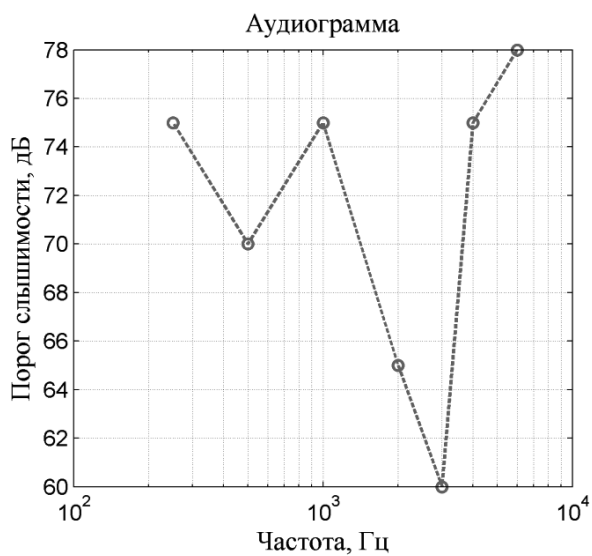
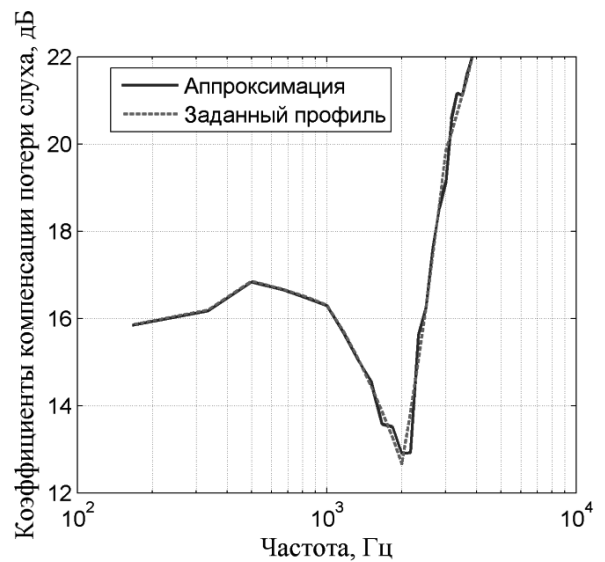
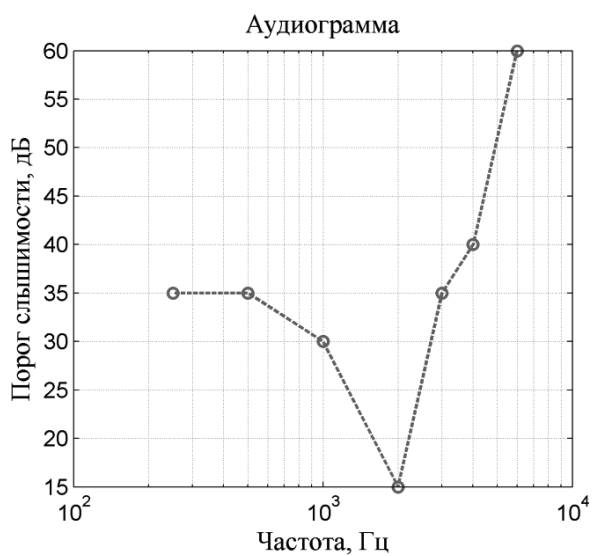
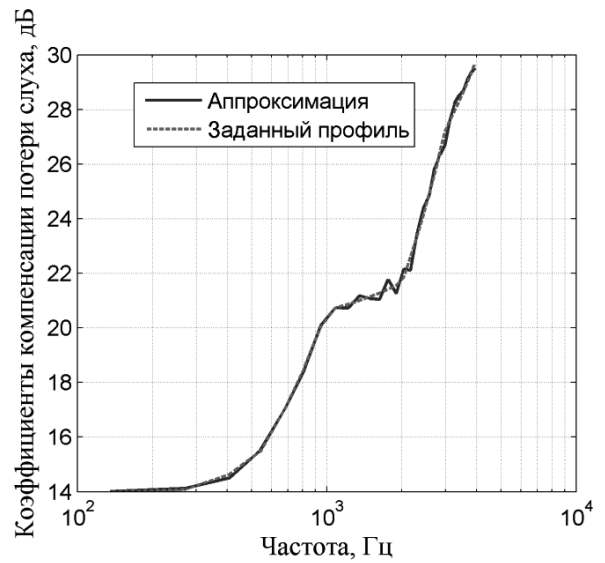
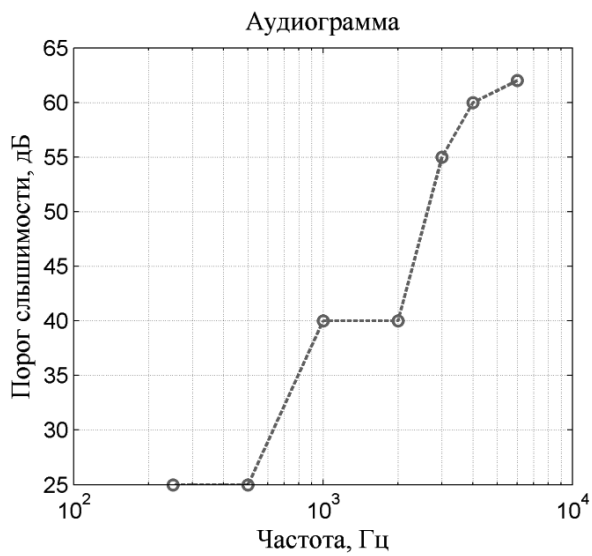


Рис. 3.9. Примеры аппроксимации профилей компенсации потери слуха при помощи МДКП (ось частот в логарифмическом масштабе)

3.5. Краткие выводы

Результаты современных исследований показывают, что перекрывающиеся преобразования могут быть эффективно применены в слуховых аппаратах. Использование перекрывающихся преобразований имеет следующие преимущества: а) малая вычислительная сложность; б) простота аппаратной реализации; в) возможность адаптации к частотному разрешению слуховой системы человека; г) малая групповая задержка. В главе приведена схема компенсации акустической обратной связи на основе ДМПП, а также схема модификации спектра сигнала на основе МДКП с деформированной частотной осью.

ГЛАВА 4

СИСТЕМАТИЧЕСКИЙ МЕТОД СИНТЕЗА БЫСТРЫХ АЛГОРИТМОВ ВЫЧИСЛЕНИЯ ДИСКРЕТНОГО КОСИНУСНОГО ПРЕОБРАЗОВАНИЯ ПРОИЗВОЛЬНОГО РАЗМЕРА

4.1. Обоснование разработки систематического метода синтеза быстрых алгоритмов ДКП

В разделе 2.1.3. показано, что эффективная реализация КМБФ связана с существованием быстрых алгоритмов вычисления дискретного косинусного преобразования (ДКП). Однако из-за того, что в таких популярных областях применения ДКП, как кодирование изображений и видео, блоки данных которых имеют размерность 4×4 , 8×8 или 16×16 , большинство разработанных быстрых алгоритмов вычисления ДКП применимы, когда формат преобразования $n = 2^k$ [92, 93]. Большинство разработанных методов применимы только к восьмиточечному ДКП. Подтверждение этому факту можно найти в [94], где указывается, что быстрые алгоритмы ДКП при $n > 8$ изучены в значительно меньшей степени, чем ДКП формата 8, которым посвящено множество работ [95–100]. Таким образом, существует потребность в разработке метода синтеза быстрых алгоритмов ДКП произвольного формата. В защиту этого утверждения можно привести следующие доводы. Быстрое развитие вычислительных платформ (таких как FPGA – Field-Programmable Gate Array) сопровождается созданием программных систем автоматической генерации структур процессоров ДКП [101, 102]. Подобные системы позволяют определить оптимальную в некотором смысле структуру процессора ДКП заданного формата для заданной вычислительной платформы. Проблема автоматического поиска структуры процессора ДКП решается как оптимизационная задача над пространством альтернативных алгоритмов вычисления ДКП. Таким образом, качество получаемого решения существенно зависит от того, ка-

кое количество альтернативных быстрых алгоритмов ДКП существует для конкретного формата преобразования.

В данной главе разрабатывается метод синтеза быстрых алгоритмов ДКП произвольного формата. В ряде случаев для одного формата преобразования метод позволяет получить несколько альтернативных быстрых алгоритмов ДКП [90].

4.2. Анализ методов синтеза быстрых алгоритмов дискретных преобразований

Для описания и синтеза быстрых алгоритмов дискретных преобразований, таких как ДКП и ДПФ, в настоящее время используются различные математические системы обозначений:

- алгебраическая система обозначений, предложенная Кули и Тьюки, используется для получения коротких рекурсивных уравнений, которые оказываются удобными при составлении программ и исследовании ошибок округления [103, 104];
- матричная система обозначений, в которой быстрый алгоритм преобразования представляется в виде произведения структурированных матриц [105];
- полиномиальные алгебры, которые позволяют перейти от матрицы преобразования к рассмотрению полиномиальных алгебр и использовать для синтеза быстрых алгоритмов математический аппарат теории групп и колец [37, 106–108].

В основе предлагаемого метода лежит подход, использующий полиномиальную алгебру, связанную с ДКП [108]. Отметим, что эффективность использования математического аппарата абстрактной алгебры для синтеза быстрых алгоритмов дискретных преобразований показана ещё в работах советских ученых [107, 109].

Полиномиальная алгебра представляет собой векторное пространство

$$\mathcal{A} = \mathbb{F}[x]/p(x). \quad (4.1)$$

Элементами алгебры является множество полиномов $\mathbb{F}[x]$, коэффициенты которых принадлежат полю \mathbb{F} , а степень меньше $n = \deg p(x)$. В полиномиальной алгебре определены операции сложения и умножения по модулю $p(x)$. В [108] показано, что синтез быстрых алгоритмов ДКП происходит в результате поэтапной декомпозиции полиномиальной алгебры (4.1), для чего требуется выполнить поэтапную факторизацию полинома $p(x)$.

Выбор поля констант \mathbb{F} в (4.1) является важной частью процесса синтеза быстрого алгоритма ДКП. Например, в [88] в качестве \mathbb{F} выбиралось поле комплексных чисел \mathbb{C} . Это удобно, поскольку над \mathbb{C} любой полином раскладывается в произведение линейных многочленов [110]. Тем не менее, выбор \mathbb{C} в качестве основного поля не дает ответа на вопрос, каким образом выполнить поэтапную факторизацию $p(x)$, необходимую для синтеза быстрого алгоритма. Для решения этой проблемы предлагается ввести в рассмотрение поле разложения полинома $p(x)$. Для выполнения последовательной факторизации $p(x)$ необходимо найти все подполя поля разложения полинома $p(x)$, для чего используется теория Галуа. В каждом подполе полином $p(x)$ имеет уникальную факторизацию, которая может быть использована для синтеза быстрого алгоритма. Поскольку коэффициенты полиномов Чебышева, которые входят в определение алгебры, отвечающей ДКП, имеют целые коэффициенты, то в качестве исходного поля предлагается выбрать поле рациональных чисел \mathbb{Q} . При выполнении поэтапной факторизации $p(x)$ поле \mathbb{Q} постепенно дополняется числами, не содержащимися в \mathbb{Q} . На последнем этапе результирующее поле является полем разложения полинома $p(x)$ [89].

Поскольку известны полиномиальные алгебры для всех восьми типов дискретных косинусных и синусных преобразований [88], то предлагаемый метод применим к синтезу быстрого алгоритма любого преобразования из данного класса.

4.3. Синтез быстрых алгоритмов дискретных преобразований с использованием полиномиальной алгебры

4.3.1. Представление алгоритма

Традиционно в цифровой обработке сигнала линейное преобразование записывается в виде

$$y_k = \sum_{0 \leq \ell < n} t_{k,\ell} s_\ell,$$

где n – формат преобразования;

$t_{k,\ell}$ – коэффициенты преобразования;

s_ℓ – входной сигнал;

y_k – выходной сигнал;

$k = 0, \dots, n - 1$ – индекс выходного сигнала.

Тем не менее, часто более удобной является векторно-матричная форма записи преобразования:

$$\mathbf{y} = T\mathbf{s},$$

где $T = [t_{k,\ell}]_{0 \leq k, \ell < n}$ – матрица преобразования;

$\mathbf{s} = [s_0, \dots, s_{n-1}]$ – входной сигнал;

$\mathbf{y} = [y_0, \dots, y_{n-1}]$ – выходной сигнал.

В этом случае быстрый алгоритм преобразования представляется в виде факторизации матрицы T в произведение слабовполненных структурированных матриц [105].

Обозначения. Все матрицы, как правило, обозначаются заглавными латинскими буквами (например A_n – квадратная матрица порядка n), а вектора – строчными латинскими буквами (полужирное начертание шрифта). Диагональные матрицы записываются как $\text{diag}(\alpha_0, \dots, \alpha_{n-1})$. Также используется оператор прямой суммы матриц

$$A \oplus B = \begin{bmatrix} A & \\ & B \end{bmatrix}.$$

Существует 8 типов дискретных косинусных преобразований [108]. Для обозначения матрицы конкретного ДКП заданного формата используется специальная нотация. Например $DCT-2_n$ обозначает матрицу ДКП второго типа формата n .

Матрицы перестановки, имеющие в точности одну 1 в строке i в позиции $f(i)$ и 0 во всех остальных, определяется как

$$P: i \mapsto f(i), \quad 0 \leq i < n.$$

Шаговая перестановка («*stride permutation*») определяется для $m|n$ как

$$L_m^n: i_2 \frac{n}{m} + i_1 \mapsto i_1 m + i_2, \quad (4.2)$$

где $0 \leq i_1 < \frac{n}{m}$; $0 \leq i_2 < m$.

Полиномы обозначаются строчными латинскими буквами, например $p(x)$, $q(x)$. Часто для удобства и экономии пространства аргумент x опускается.

4.3.2. Полиномиальное преобразование

Пусть задана полиномиальная алгебра (4.1). Предположим, что все корни $\alpha = (\alpha_0, \dots, \alpha_{n-1})$ полинома $p(x)$ попарно различны и $\alpha_k \in \mathbb{F}$. Тогда, применяя китайскую теорему об остатках (КТО), полиномиальную алгебру (4.1) можно разложить в прямую сумму одномерных подалгебр [88]:

$$\mathcal{F} : \mathbb{F}[x]/p(x) \rightarrow \bigoplus_{0 \leq k < n} \mathbb{F}[x]/(x - \alpha_k). \quad (4.3)$$

Преобразование (4.3) линейно и имеет простую интерпретацию [111]: полином степени $n - 1$ полностью определяется либо своими коэффициентами, либо списком своих значений в n различных точках. Преобразование (4.3) выполняет переход от коэффициентов полинома к его значениям в точках $\alpha = (\alpha_0, \dots, \alpha_{n-1})$. Как векторное пространство, $\mathbb{F}[x]/p(x)$ под действием \mathcal{F} раскладывается в прямую сумму одномерных подпространств $\mathbb{F}[x]/(x - \alpha_k)$. Следовательно, если зафиксировать базис $b = (p_0, \dots, p_{n-1})$ в $\mathbb{F}[x]/p(x)$ и выбрать базис $(x^0) = (1)$ в каждой подалгебре $\mathbb{F}[x]/(x - \alpha_k)$, то преобразование \mathcal{F} приобретет матричную форму:

$$\mathcal{F} = \mathcal{P}_{b,\alpha} = [p_\ell(\alpha_k)]_{0 \leq k, \ell < n}. \quad (4.4)$$

$\mathcal{P}_{b,\alpha}$ называют *полиномиальным преобразованием* алгебры $\mathbb{F}[x]/p(x)$ с базисом b . Если в каждой подалгебре $\mathbb{F}[x]/(x-\alpha_k)$ выбраны различные базисы β_k , то получающееся полиномиальное преобразование называют *масштабированным*:

$$\mathcal{F} = \text{diag}(1/\beta_1, \dots, 1/\beta_{n-1}) \cdot \mathcal{P}_{b,\alpha}. \quad (4.5)$$

Как указывалось выше, быстрый алгоритм получается путем факторизации матрицы преобразования $\mathcal{P}_{b,\alpha} = B_m B_{m-1} \dots B_1$, где B_k – слабозаполненная матрица. В терминах преобразования (4.3) быстрый алгоритм получается, если декомпозицию $\mathbb{F}[x]/p(x)$ в сумму одномерных подалгебр выполнить в несколько этапов [88].

Один из способов выполнения поэтапной декомпозиции $\mathbb{F}[x]/p(x)$ состоит в использовании факторизации $p(x) = q(x)r(x)$. Если $\deg q = k$, а $\deg r = m$, то

$$\begin{aligned} & \mathbb{F}[x]/p(x) \\ \rightarrow & \mathbb{F}[x]/q(x) \oplus \mathbb{F}[x]/r(x) \end{aligned} \quad (4.6)$$

$$\rightarrow \bigoplus_{0 \leq i < k} \mathbb{F}[x]/(x - \beta_i) \oplus \bigoplus_{0 \leq j < m} \mathbb{F}[x]/(x - \gamma_j) \quad (4.7)$$

$$\rightarrow \bigoplus_{0 \leq i < n} \mathbb{F}[x]/(x - \alpha_i), \quad (4.8)$$

где β_i – нули полинома $q(x)$;

γ_j – нули полинома $r(x)$.

Если в подалгебре $\mathbb{F}[x]/q(x)$ выбрать базис c , а в $\mathbb{F}[x]/r(x)$ – базис d , тогда поэтапную декомпозицию (4.6)–(4.8) можно записать в виде произведения матриц:

$$\mathcal{P}_{b,\alpha} = P(\mathcal{P}_{c,\beta} \oplus \mathcal{P}_{d,\gamma})B. \quad (4.9)$$

Матрица B отображает базис b в конкатенацию базисов (c, d) и соответствует выражению (4.6). Преобразованию (4.7), выполняющему декомпозицию $\mathbb{F}[x]/q(x)$ и $\mathbb{F}[x]/r(x)$ с использованием КТО, соответствует прямая сумма матриц $\mathcal{P}_{c,\beta}$ и $\mathcal{P}_{d,\gamma}$. На этапе (4.8) происходит перестановка одномерных алгебр. В выражении (4.9) этому шагу отвечает матрица P , выполняющая отображение $(\beta, \gamma) \mapsto \alpha$. Если B – слабозаполненная матрица, тогда (4.9) представляет собой быстрый алгоритм, поскольку оставшиеся матрицы-сомножители являются слабозаполненными по определению.

Описанная выше процедура применяется в дальнейшем для синтеза быстрых алгоритмов ДКП.

4.4. Алгебраический метод синтеза быстрых алгоритмов ДКП

4.4.1. Описание метода

Предлагаемый алгебраический метод синтеза быстрых алгоритмов ДКП состоит в выполнении следующих шагов.

Шаг 1. Определение полиномиальной алгебры $\mathbb{Q}[x]/p(x)$ (и базиса в ней), отвечающей заданному типу ДКП.

Шаг 2. Получение всех подполей \mathbb{L}_i поля разложения \mathbb{E} полинома $p(x)$ с использованием теории Галуа.

Шаг 3. Получение поэтапной факторизации полинома $p(x)$, отвечающей башне вложенных подполей $\mathbb{Q} = \mathbb{L}_0 \subset \mathbb{L}_1 \subset \dots \subset \mathbb{L}_r = \mathbb{E}$.

Шаг 4. Синтез быстрого алгоритма ДКП с использованием факторизации, полученной на шаге 3, и выражений (4.6)—(4.8).

Полиномиальные алгебры, отвечающие всем 8 типам дискретных косинусных и синусных преобразований, можно найти в [88, 108]. Необходимые для понимания метода сведения из теории Галуа приведены в приложении А.

4.4.2. Синтез быстрого алгоритма ДКП-4 формата 7

Для конкретизации предлагаемого метода ниже рассматривается задача синтеза быстрого алгоритма ДКП четвертого типа. Традиционные методы синтеза быстрых алгоритмов ДКП, как правило, применимы в случае, когда формат преобразования равен составному числу [88]. Чтобы показать преимущество предлагаемого метода, выполним синтез быстрого алгоритма ДКП-4, формат которого равен простому числу 7.

Шаг 1. Рассмотрим полиномиальную алгебру, отвечающую ДКП-4 [108]:

$$\mathcal{A} = \mathbb{Q}[x]/2T_n(x), \quad b = (V_0, \dots, V_{n-1}), \quad (4.10)$$

где T и V – полиномы Чебышева первого и третьего рода соответственно. Краткое описание полиномов Чебышева 1–4 рода приведено в приложении Б.

Поскольку корни $2T_n(x)$ равны $\alpha_k = \cos(k + \frac{1}{2})\frac{\pi}{n}$, $0 \leq k < n$, то в соответствии с (4.4) полиномиальное преобразование для (4.10) определяется как

$$\mathcal{P}_{\alpha,b} = [V_\ell(\alpha_k)]_{0 \leq k, \ell < n} = \left[\frac{\cos(k + \frac{1}{2})(\ell + \frac{1}{2})\frac{\pi}{n}}{\cos(k + \frac{1}{2})\frac{\pi}{2n}} \right]. \quad (4.11)$$

Если (4.11) умножить слева на масштабирующую диагональную матрицу

$$D_n^{(C4)} = \text{diag}_{0 \leq k < n} \left(\cos(k + \frac{1}{2})\frac{\pi}{2n} \right), \quad (4.12)$$

то в результате получим матрицу ДКП-4:

$$\text{DCT-4}_n = \left[\cos(k + \frac{1}{2})(\ell + \frac{1}{2})\frac{\pi}{n} \right]_{0 \leq k, \ell < n}. \quad (4.13)$$

Таким образом, выражения (4.11)–(4.13) демонстрируют, что ДКП-4 представляет собой масштабированное полиномиальное преобразование вида (4.5).

Полиномиальное преобразование, соответствующее дискретному тригонометрическому преобразованию (ДТТ), обозначается как $\overline{\text{ДТТ}}$. Например, $\overline{\text{DCT-4}_n}$ соответствует матрице (4.11).

Шаг 2. Полиномиальная алгебра, соответствующая 7-точечному ДКП-4 имеет вид

$$\mathcal{A} = \mathbb{Q}[x]/2T_7(x), \quad b = (V_0, \dots, V_6). \quad (4.14)$$

Корни $2T_7(x)$ равны $\alpha_k = \cos(k + \frac{1}{2})\frac{\pi}{14}$, $k = 0, 1, \dots, 6$. Над полем \mathbb{Q} данный полином имеет факторизацию

$$2T_7(x) = 2xP_6(x), \quad (4.15)$$

поскольку корень $\alpha_3 = 0$. Факторизация (4.15) будет использована при получении быстрого алгоритма.

Полином $P_6(x)$ неприводим над полем \mathbb{Q} . Его полем разложения служит $\mathbb{Q}_{\cos(\pi/14)}$, поскольку все корни $P_6(x)$ могут быть выражены через примитивный элемент $\theta = \cos(\frac{\pi}{14})$ следующим образом:

$$\alpha_k = T_{1+2k}(\theta), \quad k = 0, 1, 2, 4, 5, 6. \quad (4.16)$$

Согласно описанному методу синтеза быстрых алгоритмов ДКП необходимо определить подполя поля $\mathbb{Q}_{\cos(\pi/14)}$, в которых $P_6(x)$ может быть разложен на множители. Для этого определим группу Галуа $\text{Gal}(P_6)$. Воспользуемся методом, описанным в [112], для отыскания всех автоморфизмов поля $\mathbb{Q}_{\cos(\pi/14)}$. Введем следующие обозначения $\theta_0 = \alpha_0$, $\theta_1 = \alpha_1$, $\theta_2 = \alpha_2$, $\theta_3 = \alpha_4$, $\theta_4 = \alpha_5$, $\theta_5 = \alpha_6$. Поскольку элементы $\theta_0, \dots, \theta_5$ являются числами, сопряженными с θ (и могут быть выражены через θ , см. (4.16)), тогда подстановки

$$\theta \mapsto \theta_k, \quad k = 0, 1 \dots 5, \quad (4.17)$$

исчерпывают всю группу Галуа $\text{Gal}(P_6)$. Таким образом, каждый элемент группы Галуа переводит систему корней $P_6(x)$ в себя. Подстановки (4.17) удобно представлять матрицами перестановок $\{\sigma_0, \dots, \sigma_5\}$, которые имеют разряженную структуру (рис. 4.1).

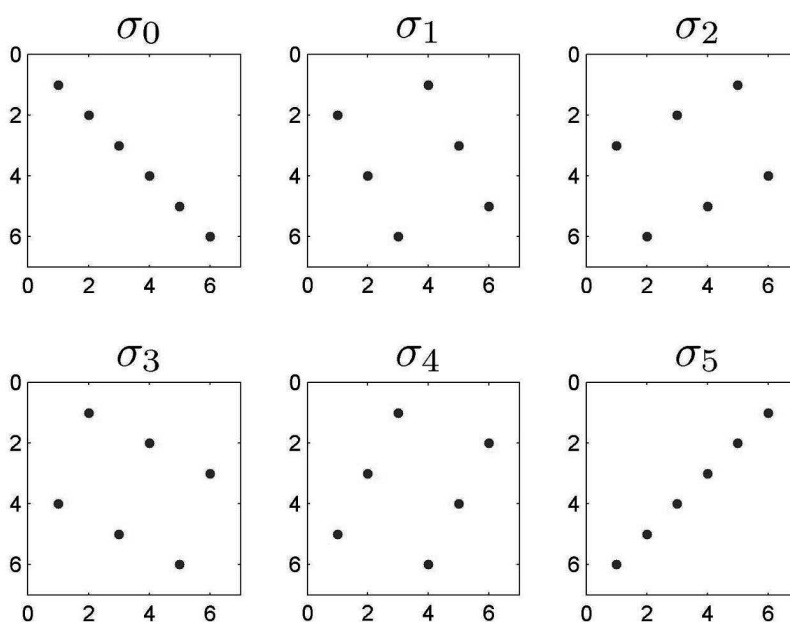


Рис. 4.1. Группа Галуа $\text{Gal}(P_6)$, представленная матрицами перестановок $\{\sigma_0, \dots, \sigma_5\}$. Кружками обозначены положения единиц в матрице.

Таким образом, установлено, что $\text{Gal}(P_6)$ представляет собой циклическую группу шестого порядка (\mathbb{Z}_6). Как известно, \mathbb{Z}_6 имеет две подгруппы – циклическую подгруппу порядка 2: $\mathbb{Z}_2 = \{\sigma_0, \sigma_5\}$, и циклическую подгруппу порядка 3: $\mathbb{Z}_3 = \{\sigma_0, \sigma_1, \sigma_2\}$. На рис. 4.2, а группа $\text{Gal}(P_6)$ показана в виде решетки подгрупп. Согласно соответствию Галуа должно

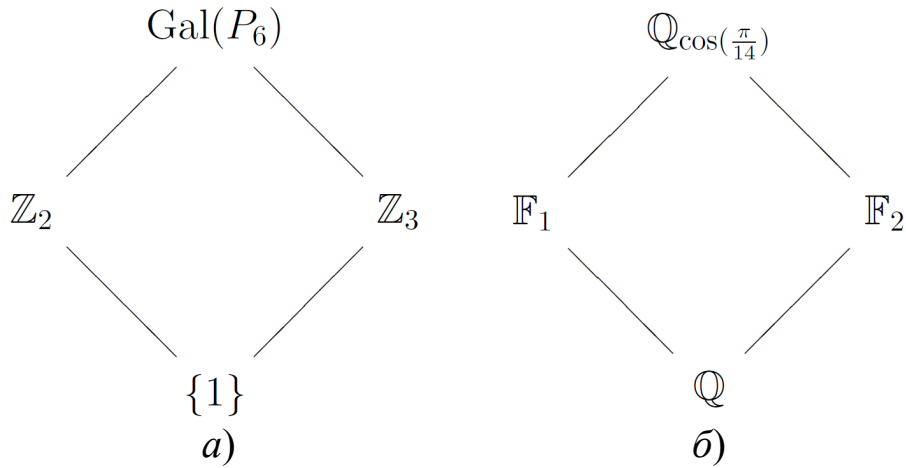


Рис. 4.2. а) подгруппы группы Галуа $\text{Gal}(P_6)$; б) подполя поля разложения $P_6(x)$

существовать аналогичная решетка подполей поля $\mathbb{Q}_{\cos(\pi/14)}$ (рис. 4.2, б).

Возникает следующий вопрос: как найти подполе \mathbb{F} , соответствующее некой подгруппе H группы Галуа? В [112] предлагается следующее решение. Если задана подгруппа $H = \{\sigma_0, \dots, \sigma_{h-1}\}$, то необходимо составить произведение

$$(x - \sigma_0\theta)(x - \sigma_1\theta) \dots (x - \sigma_{h-1}\theta), \quad (4.18)$$

коэффициенты этого полинома согласно основной теореме теории Галуа должны принадлежать полю \mathbb{F} и даже порождать его.

Применим данный способ для нахождения подполей \mathbb{F}_1 и \mathbb{F}_2 (рис. 4.2, б). Рассмотрим полином

$$(x - \sigma_0\theta)(x - \sigma_5\theta) = x^2 - \frac{1}{2} \cos \frac{6\pi}{7},$$

коэффициенты которого принадлежат полю $\mathbb{Q}_{\cos \frac{\pi}{7}}$. Это и есть искомое подполе \mathbb{F}_1 . Подполе \mathbb{F}_2 определяется аналогичным образом, если рассмотреть полином

$$\begin{aligned} (x - \sigma_0\theta)(x - \sigma_1\theta)(x - \sigma_3\theta) &= \\ &= x^3 + x^2 \left(\cos \frac{\pi}{14} + \cos \frac{3\pi}{14} + \cos \frac{9\pi}{14} \right) - \frac{1}{4} \left(\cos \frac{\pi}{14} + \cos \frac{3\pi}{14} + \cos \frac{9\pi}{14} \right), \end{aligned}$$

коэффициенты которого лежат в поле $\mathbb{Q}_{\cos \frac{\pi}{14} + \cos \frac{3\pi}{14} + \cos \frac{9\pi}{14}} \cong \mathbb{Q}_{\sqrt{7}}$, соответствующему подполю \mathbb{F}_2 .

Шаг 3. Используя две башни полей на рис. 4.2, б получаются два

различных способа поэтапной факторизации полинома $2T_7(x)$:

$$\begin{aligned}
 2T_7(x) &= \underbrace{2xP_6(x)}_{\mathbb{Q}} = \\
 &= \underbrace{2^4 x (T_2(x) - \cos \frac{6\pi}{7})(T_2(x) - \cos \frac{4\pi}{7})(T_2(x) - \cos \frac{2\pi}{7})}_{\mathbb{Q}_{\cos \frac{\pi}{7}}} = \\
 &= \underbrace{2^7 \prod_{i=0}^6 (x - \alpha_i)}_{\mathbb{Q}_{\cos \frac{\pi}{14}}} \quad (4.19)
 \end{aligned}$$

$$\begin{aligned}
 2T_7(x) &= \underbrace{2xP_6(x)}_{\mathbb{Q}} = \\
 &= \underbrace{2x(T_3(x) - 2\sqrt{7}T_2(x) + 6T_1(x) - \sqrt{7})(T_3(x) + 2\sqrt{7}T_2(x) + 6T_1(x) + \sqrt{7})}_{\mathbb{Q}_{\sqrt{7}}} = \\
 &= \underbrace{2^7 \prod_{i=0}^6 (x - \alpha_i)}_{\mathbb{Q}_{\cos \frac{\pi}{14}}} \quad (4.20)
 \end{aligned}$$

Шаг 4. На основе факторизаций (4.19)–(4.20) и метода синтеза быстрых алгоритмов ДКП, описанного в начале данного раздела, получены два быстрых алгоритма 7-точечного ДКП-4. Детали процедуры получения быстрого алгоритма по известной факторизации полинома можно найти в работах [88, 89].

На рис. 4.3 показана граф-схема быстрого алгоритма 7-точечного ДКП-4, полученная с использованием факторизации (4.19). Штрихпунктирными линиями обозначены сигналы, которые в блоках суммирования берутся с отрицательным знаком. Пунктирной линией обозначены внутренние сигналы, которые умножаются на константу (соответствующая константа приведена рядом). На рис. 4.3 не показана операция масштабирования, которая соответствует умножению вектора выходных данных граф-схемы алгоритма на диагональную матрицу $D_7^{(C4)}$ (см. выражение (4.12)). Без учета масштабирующих множителей быстрый алгоритм 7-точечного ДКП-4 (рис. 4.3) требует 15 операций умножения и 37 сложений.

4.4.3. Рекурсивный быстрый алгоритм ДКП-4

Рассмотрим вопрос синтеза быстрого алгоритма ДКП-4 формата 2^k с использованием алгебраического метода, описанного в пункте 4.4.1..

Полиномиальная алгебра, отвечающая ДКП-4 формата 2^k , имеет вид [108]

$$\mathcal{A} = \mathbb{Q}[x]/2T_{2^k}(x), \quad b = (V_0, \dots, V_{2^k-1}).$$

Можно показать, что группа Галуа полинома $2T_{2^k}(x)$ является циклической группой, имеющей порядок 2^k :

$$\text{Gal}(2T_{2^k}) \cong \mathbb{Z}_{2^k}. \quad (4.22)$$

Последнее выражение приводит к следующему ряду нормальных вложенных подгрупп:

$$\text{Gal}(2T_{2^k}) \cong \mathbb{Z}_{2^k} \supset \mathbb{Z}_{2^{k-1}} \supset \dots \supset \mathbb{Z}_2 \supset \{1\}, \quad (4.23)$$

и соответствующей башне полей

$$\mathbb{Q} \subset \mathbb{Q}_{\sqrt{2}} \subset \dots \subset \mathbb{Q}_{\underbrace{\sqrt{2+\sqrt{2+\dots+\sqrt{2}}}}_k}. \quad (4.24)$$

Используя (4.22)–(4.24), факторизацию полинома $2T_{2^k}(x)$ можно записать в рекурсивном виде. Для получения общей формулы вначале рассмотрим специальный случай $k = 2$. Нули $2T_4(x)$ равны $\alpha_\ell = \cos(\ell + \frac{1}{2})\frac{\pi}{4}$, $\ell = 0, \dots, 3$:

$$\begin{aligned} \alpha_0 &= \cos \frac{\pi}{8} = -\alpha_3 = -\cos \frac{7\pi}{8} = \frac{1}{2}\sqrt{2 + \sqrt{2}}; \\ \alpha_1 &= \cos \frac{3\pi}{8} = -\alpha_2 = -\cos \frac{5\pi}{8} = \frac{1}{2}\sqrt{2 - \sqrt{2}}. \end{aligned}$$

Поле разложения $2T_4(x)$ является $\mathbb{Q}_{\sqrt{2+\sqrt{2}}}$. Данное поле порождается любым из корней $2T_4(x)$. Выберем в качестве примитивного элемента $\theta = \alpha_0$, тогда группа Галуа $2T_4(x)$ исчерпывается следующими подстановками:

$$\sigma_k : \theta \mapsto \alpha_k, \quad k = 0, \dots, 3. \quad (4.25)$$

Каждый корень $2T_4(x)$ может быть выражен через примитивный элемент θ : $\alpha_k = T_{2k+1}(\theta)$. Используя этот факт и выражение (4.25), можно определить таблицу Кэли группы $\text{Gal}(2T_4)$ – табл. 4.1.

По табл. 4.1 легко определить, что

- $H \in \{\sigma_0, \sigma_3\}$ является подгруппой $\text{Gal}(2T_4)$;
- $\text{Gal}(2T_4) \cong \mathbb{Z}_4$;
- $H \cong \mathbb{Z}_2$.

Подгруппа H определяет промежуточное подполе

$$\mathbb{Q}_H = \{\eta \in \mathbb{Q}_{\sqrt{2+\sqrt{2}}} \mid g(\eta) = \eta \ \forall g \in H\}.$$

Очевидно, что $\mathbb{Q}_H \cong \mathbb{Q}_{\sqrt{2}}$, используя соответствие Галуа, получаем

$$\mathbb{Z}_4 \supset \mathbb{Z}_2 \supset \{1\} \Rightarrow \mathbb{Q} \subset \mathbb{Q}_{\sqrt{2}} \subset \mathbb{Q}_{\sqrt{2+\sqrt{2}}}.$$

Таблица 4.1. Таблица Кэли группы Галуа Gal_{2T_4}

\cdot	σ_0	σ_1	σ_2	σ_3
σ_0	σ_0	σ_1	σ_2	σ_3
σ_1	σ_1	σ_3	σ_0	σ_2
σ_2	σ_2	σ_0	σ_3	σ_1
σ_3	σ_3	σ_2	σ_1	σ_0

Факторизация $2T_4(x)$, получающаяся с использованием приведенной башни полей, имеет следующий вид

$$\begin{aligned} \mathbb{Q}: & 2T_4(x); \\ \mathbb{Q}_{\sqrt{2}}: & (2T_2(x) - \sqrt{2})(2T_2(x) + \sqrt{2}); \\ \mathbb{Q}_{\sqrt{2+\sqrt{2}}}: & (2T_1(x) + \sqrt{2+\sqrt{2}})(2T_1(x) - \sqrt{2+\sqrt{2}}) \\ & (2T_1(x) + \sqrt{2-\sqrt{2}})(2T_1(x) - \sqrt{2-\sqrt{2}}). \end{aligned}$$

Из приведенного частного случая путем индукции по k можно вывести, что T_{2^k} раскладывается на множители вида $(2T_n - 2 \cos r\pi)$, и данный процесс можно представить в виде следующей рекурсивной формулы:

$$2T_{2^k}(x) - 2 \cos r\pi = \left(2T_{2^{k-1}}(x) - 2 \cos \frac{r\pi}{2}\right) \left(2T_{2^{k-1}}(x) - 2 \cos \pi\left(1 - \frac{r}{2}\right)\right). \quad (4.26)$$

где $r \in [0; 1]$ – параметр.

Выражение (4.26) можно доказать, используя аналитическую запись полинома $T_{2^k}(x)$. Факторизация $2T_{2^k}(x)$ определяется формулой (4.26), если

выбрать $r = 1/2$. В силу того, что $2 \cos \frac{r\pi}{2} = \sqrt{2 + 2 \cos r\pi}$, левая и правая части выражения (4.26) определены над различными полями: $\mathbb{Q}_{2 \cos r\pi}$ и $\mathbb{Q}_{\sqrt{2+2 \cos r\pi}}$ соответственно. Выражение (4.26) определяет последовательный переход от поля \mathbb{Q} к полю разложения полинома $2T_{2^k}(x)$ в соответствии с (4.24).

Факторизация (4.26) позволяет получить быстрый алгоритм для более общего класса асимметричных (*skew*) ДКП-4 [88], которые описываются алгеброй

$$\mathcal{A} = \mathbb{Q}_{\cos r\pi}[x]/(2T_{2^k}(x) - 2 \cos r\pi), \quad b = (V_0, \dots, V_{2^k-1}). \quad (4.27)$$

Матрица асимметричного ДКП-4 обозначается как $\text{DCT-4}_n(r)$. Обычное ДКП-4 является специальным случаем асимметричного при $r = 1/2$.

Используя факторизацию (4.26), алгебру (4.27) можно разложить в прямую сумму подалгебр:

$$\begin{aligned} & \mathbb{Q}_{\cos r\pi}[x]/(2T_{2^k}(x) - 2 \cos r\pi) \\ \rightarrow & \mathbb{Q}_{\cos \frac{r\pi}{2}}[x]/(2T_{2^{k-1}}(x) - 2 \cos \frac{r\pi}{2}) \oplus \\ & \mathbb{Q}_{\cos \frac{r\pi}{2}}[x]/(2T_{2^{k-1}}(x) - 2 \cos \pi(1 - \frac{r}{2})). \end{aligned} \quad (4.28)$$

Учитывая (4.6)—(4.9) и выбирая в подалгебрах базис $c = d = (V_0, \dots, V_{2^k/2-1})$, можно получить следующий быстрый алгоритм для $\overline{\text{DCT-4}}_{2^k}(r)$:

$$\overline{\text{DCT-4}}_{2^k}(r) = P \cdot (\overline{\text{DCT-4}}_{2^{k-1}}(\frac{r}{2}) \oplus \overline{\text{DCT-4}}_{2^{k-1}}(1 - \frac{r}{2})) \cdot B_{2^k}^{(C4)}, \quad (4.29)$$

где P – матрица перестановки вида

$$P = \begin{bmatrix} 1 & & & & & \\ & I_2 & & & & \\ & & I_2 & & & \\ & & & \ddots & & \\ & & & & \ddots & \\ & & & & & 1 \\ & & & & & & I_2 \end{bmatrix},$$

а $B_{2^k}^{(C4)}$ – матрица перехода от базиса b к конкатенации базисов (c, d) . Полиномы $V_\ell \in b$, $0 \leq \ell < m = 2^{k-1}$ содержатся непосредственно в базисах c и d . Поэтому первые 2^{k-1} столбцов матрицы имеют вид

$$B_{2^k}^{(C4)} = \begin{bmatrix} I_{2^{k-1}} & * \\ I_{2^{k-1}} & * \end{bmatrix}.$$

Оставшиеся элементы матрицы определяются, используя следующие выражения, которые показывают куда отображаются полиномы $V_{m+\ell} \in b$ в новом базисе:

$$\begin{aligned} V_{m+\ell} &\equiv -V_{m-\ell-1} + 2 \cos \frac{r\pi}{2} V_\ell \pmod{2T_m - 2 \cos \frac{r\pi}{2}}; \\ V_{m+\ell} &\equiv -V_{m-\ell-1} - 2 \cos \frac{r\pi}{2} V_\ell \pmod{2T_m - 2 \cos \pi(1 - \frac{r}{2})}. \end{aligned} \quad (4.30)$$

Данные уравнения выведены исходя из известных соотношений для полиномов Чебышева: $2T_m = V_m + V_{m-1}$ и $V_m = 2xV_{m-1} - V_{m-2}$. Учитывая (4.30),

$$\begin{aligned} B_{2^k}^{(C4)}(r) &= \begin{bmatrix} I_{2^{k-1}} & (2 \cos \frac{r\pi}{2} I_{2^{k-1}} - J_{2^{k-1}}) \\ I_{2^{k-1}} & (-2 \cos \frac{r\pi}{2} I_{2^{k-1}} - J_{2^{k-1}}) \end{bmatrix} = \\ &= \begin{bmatrix} I_{2^{k-1}} & I_{2^{k-1}} \\ I_{2^{k-1}} & -I_{2^{k-1}} \end{bmatrix} \cdot \begin{bmatrix} I_{2^{k-1}} & -J_{2^{k-1}} \\ & 2 \cos \frac{r\pi}{2} I_{2^{k-1}} \end{bmatrix}. \end{aligned} \quad (4.31)$$

Заметим, что декомпозиция (4.28) требует расширения основного поля констант $\mathbb{Q}_{\cos r\pi}$ до $\mathbb{Q}_{\cos \frac{r\pi}{2}}$. Элементы нового поля констант появляются в матрице смены базиса $B_{2^k}^{(C4)}(r)$.

В итоге быстрый алгоритм асимметричного ДКП-4 получается, если (4.29) домножить слева на диагональную матрицу $D_{2^k}^{(C4)}$:

$$\text{DCT-4}_{2^k}(r) = D_{2^k}^{(C4)} \overline{\text{DCT-4}_{2^k}(r)}.$$

Ниже приводится `Matlab`-реализация предложенного рекурсивного быстрого алгоритма (4.29).

```
function [y] = dct4(r, x)
%DCT4 Computation of skew DCT4(r).
m = length(x)/2;
Im = eye(m);          Jm = rot90(Im);
a = 2*cos(pi*r/2);    Zm = zeros(m,m);
B1 = [Im -Jm; Zm a*Im];
B2 = [Im Im; Im -Im];
x = B2*B1*x;
P = eye(2);
if m~=1
```



```

x1 = dct4(r/2, x(1:m));
x2 = dct4(1-r/2, x(m+1:2*m));
x = [x1; x2];
P = zeros(2*m, 2*m);
P(1,1) = 1; P(2*m,m) = 1;
row = 2; col = m+1; sign = -1;
for i=1:m-1
    P(row,col) = 1;
    row = row+1; col = col+1;
    P(row,col) = 1;
    row = row+1; col = col+sign*m;
    sign= sign*(-1);
end
end
y = P*x;
end

```

4.4.4. Рекурсивный быстрый алгоритм ДКП-2

Рассмотрим вопрос синтеза быстрого алгоритма ДКП-2. Полиномиальная алгебра, отвечающая данному преобразованию, имеет вид

$$\mathcal{A} = \mathbb{Q}[x]/(x-1)U_{n-1}(x), \quad b = (V_0, \dots, V_{n-1}). \quad (4.32)$$

где n – формат преобразования;

$U_{n-1}(x)$ – полином Чебышева второго рода.

Поскольку нули полинома $U_n(x)$ равны $\alpha_k = \cos \frac{(k+1)\pi}{n+1}$, $0 \leq k < n$, то полиномиальное преобразование для (4.32) принимает вид

$$\mathcal{P}_{\alpha,b} = [V_\ell(\alpha_k)]_{0 \leq k, \ell < n} = \overline{\text{DCT-2}}_n = \left[\frac{\cos k(\ell + \frac{1}{2})\frac{\pi}{n}}{\cos \frac{k\pi}{2n}} \right]. \quad (4.33)$$

Для получения матрицы DCT-2 необходимо (4.33) умножить слева на масштабирующую диагональную матрицу

$$D_n^{(C2)} = \text{diag}_{0 \leq k < n} \left(\cos \frac{k\pi}{2n} \right).$$

Известно [93], что вычисление ДКП-2 формата $2n$ можно свести к вычислению ДКП-2 и ДКП-4 формата n . С точки зрения алгебраического метода синтеза быстрых алгоритмов [88] это объясняется наличием факторизации

$$U_{2n-1}(x) = U_{n-1}(x) \cdot 2T_n(x),$$

благодаря которой алгебра $\mathbb{Q}[x]/(x-1)U_{2n-1}(x)$ (с базисом $b = (V_0, \dots, V_{2n-1})$) может быть разложена в прямую сумму подалгебр

$$\begin{aligned} & \mathbb{Q}[x]/(x-1)U_{2n-1}(x) \\ \rightarrow & \mathbb{Q}[x]/(x-1)U_{n-1}(x) \oplus \mathbb{Q}[x]/2T_n(x). \end{aligned} \quad (4.34)$$

В соответствии с (4.6)–(4.8) выражение (4.34) ведет к следующему быстрому алгоритму [108]:

$$\overline{\text{DCT-2}}_{2n} = L_n^{2n} (\overline{\text{DCT-2}}_n \oplus \overline{\text{DCT-4}}_n) B_{2n}^{(C2)}, \quad (4.35)$$

где L_n^{2n} – матрица шаговой перестановки (см. (4.2));

$B_{2n}^{(C2)}$ – матрица смены базиса.

Матрица $B_{2n}^{(C2)}$ отображает базис b в конкатенацию базисов (c, d) , где $c = d = (V_0, \dots, V_{n-1})$ – базисы подалгебр, находящихся в правой части выражения (4.34). Первые n столбцов $B_{2n}^{(C2)}$ определяются как

$$B_{2n}^{(C2)} = \begin{bmatrix} I_n & * \\ I_n & * \end{bmatrix},$$

поскольку $V_\ell \in b$, $0 \leq \ell < n$, изначально содержатся в c и d . Оставшиеся элементы матрицы определяются следующими выражениями:

$$\begin{aligned} V_{n+\ell} &\equiv V_{n-\ell-1} \pmod{(x-1)U_n}; \\ V_{n+\ell} &\equiv -V_{n-\ell-1} \pmod{2T_n}, \end{aligned}$$

откуда

$$B_{2n}^{(C2)} = \begin{bmatrix} I_n & J_n \\ I_n & -J_n \end{bmatrix}.$$

Заметим, что декомпозиция (4.34) не требует расширения основного поля констант \mathbb{Q} . Это приводит к тому, что матрица $B_{2n}^{(C2)}$ не содержит нетривиальных множителей.

В случае, когда формат преобразования равен степени числа 2, совместное использование факторизаций (4.35) и (4.29) позволяет получить быстрый рекурсивный алгоритм DCT-2_{2^k} [91]. Основной операцией алгоритма является умножение на матрицу $B_n^{(C4)}(r)$. Все нетривиальные множители сосредоточены в этой матрице, которая по сути играет ту же роль, что и матрица поворачивающих множителей в алгоритмах быстрого преобразования Фурье Кули–Тьюки.

Быстрый алгоритм 16-точечного ДКП-2. В качестве примера предлагаемого подхода рассмотрим синтез быстрого алгоритма 16-точечного ДКП-2. Вначале матрица преобразования представляется в виде произведения:

$$\text{DCT-2}_{16} = D_{16}^{(C2)} \cdot \overline{\text{DCT-2}_{16}}.$$

Далее факторизации (4.35) и (4.29) применяются рекурсивно для получения быстрого алгоритма (рис. 4.4). На рисунке не показана операция масштабирования, которая соответствует умножению вектора выходных данных граф-схемы алгоритма на диагональную матрицу $D_{16}^{(C2)}$. Базовый вычислительный блок алгоритма, выполняющий умножение на матрицу (4.31), приведен на рис. 4.5. Все операции внутри базового блока выполняются над m компонентными векторами. Реализация базового блока требует $3m$ операций сложения и m операций умножения. Представленный быстрый алгоритм 16-точечного ДКП-2 имеет сложность в 32 умножения и 81 сложение. Однако только 17 нетривиальных умножений составляют вычислительное ядро алгоритма. Оставшиеся 15 – масштабирующие множители, которые могут быть учтены при пост-обработке. Кроме того, алгоритм на рис. 4.4 включает вычисление масштабированного 8-точечного ДКП-2, которое требует только 5 операций умножения. Это такое же число операций, как и в алгоритме Араи [96], который используется в стандарте JPEG. Однако алгоритм Араи получен путем манипуляций с элементами матрицы ДКП-2, в то время как предлагаемый подход имеет более общий характер. В табл. 4.2 приводится сравнение полученного быстрого алгоритма 16-точечного ДКП-2 с известными алгоритмами.

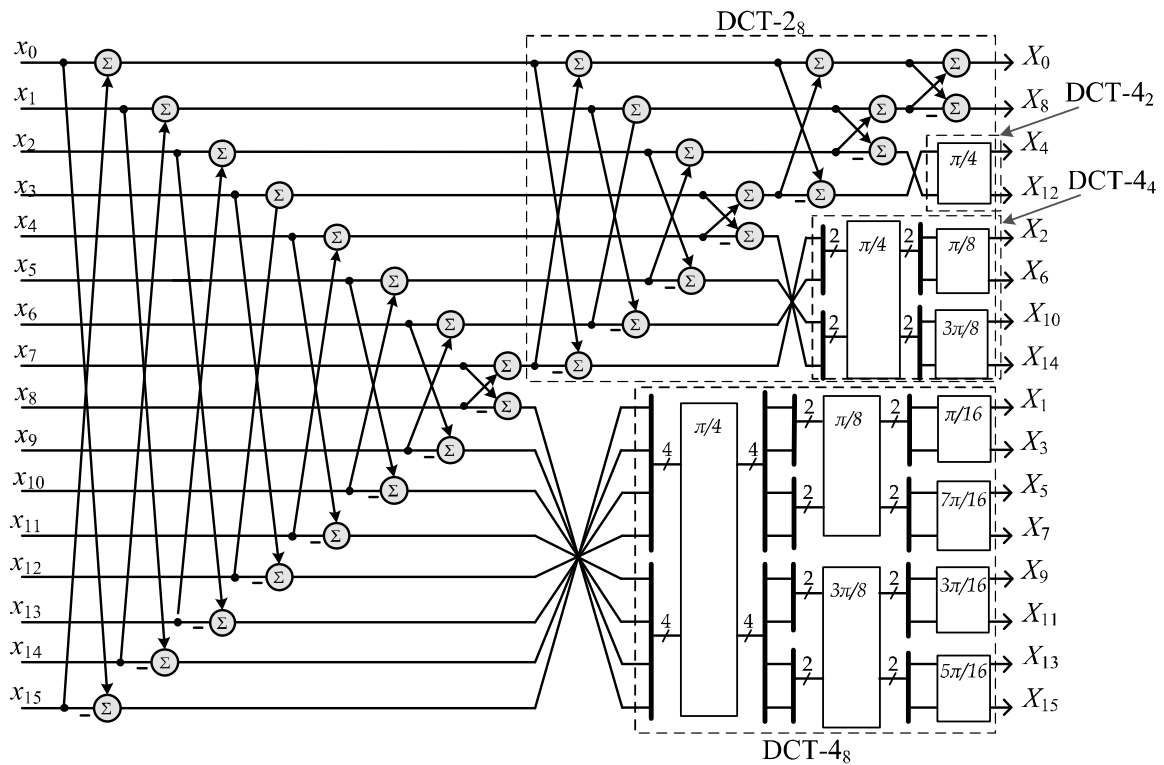


Рис. 4.4. Граф-схема быстрого алгоритма 16-точечного ДКП-2

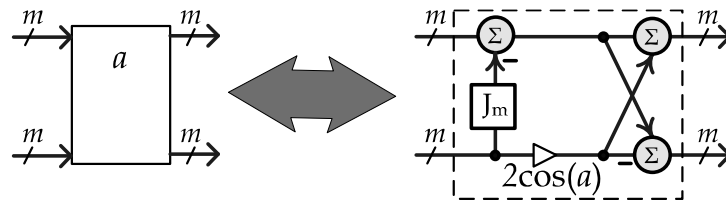


Рис. 4.5. Базовый вычислительный блок для быстрых алгоритмов ДКП-2/ДКП-4

Таблица 4.2. Число операций для вычисления 16-точечного ДКП-2

Автор (ссылка)	Chen [95]	Wang [113]	Lee [114]	Loeffler [97]	Hou [115]	Предложен- ный алго- ритм
Число умножений	44	35	32	31	32	32
Число сложений	74	83	81	81	81	81

4.5. Основные результаты и выводы

1. Предложен алгебраический метод синтеза быстрых алгоритмов ДКП произвольного формата. В методе используется представление ДКП как матрицы декомпозиции определенной полиномиальной алгебры $\mathbb{F}[x]/p(x)$. Быстрый алгоритм получается в результате выполнения поэтапной декомпозиции данной алгебры. В свою очередь декомпозиция тесно связана с выполнением факторизации полинома $p(x)$. Для решения этой задачи используется математический аппарат теории Галуа, при помощи которого находятся все подполя поля разложения полинома $p(x)$. В данных подполях $p(x)$ имеет факторизацию, которая используется для получения быстрого алгоритма. В качестве практического применения получены два варианта быстрого алгоритма 7-точечного ДКП-4.

2. Предложенный метод также применен для синтеза быстрых алгоритмов ДКП-4/ДКП-2, когда формат преобразования равен степени числа 2. Полученные в результате алгоритмы обладают регулярной структурой, а их вычислительное ядро обладает малой мультипликативной сложностью.

ГЛАВА 5

АЛГОРИТМЫ ОБРАБОТКИ СИГНАЛА В СЛУХОВОМ АППАРАТЕ НА ОСНОВЕ НЕРАВНОПОЛОСНОГО БАНКА ФИЛЬТРОВ

5.1. Модификация спектра сигнала в слуховом аппарате

В главе рассматриваются алгоритмы обработки сигнала в слуховом аппарате, в котором в качестве средства частотного анализа выступает неравнополосный банк фильтров. Ранее уже обосновывалось требование малой групповой задержки слухового аппарата, связанное с эффектом эха. В данной главе для уменьшения групповой задержки сигнала в слуховом аппарате предлагается использовать неравнополосный косинусно-модулированный банк фильтров с малой длинной фильтра-прототипа, который можно рассматривать как модифицированное ДКП с деформированной частотной осью.

Перекрывающееся преобразование (*lapped transform*), особенное МДКП широко применяется в высококачественном сжатии аудио сигналов [54]. В этом приложении требуются хорошая частотная избирательность и отсутствие избыточности в субканалах, поэтому для кодирования используют длинные временные окна (до 2048 коэффициентов) и критическую децимацию субканальных сигналов. Предполагается, что для слуховых аппаратов можно получить более простое перекрывающееся преобразование с небольшим числом субканалов, в которых отсутствует децимация. Это позволит избежать проблем с наложением спектра в следствии передискретизации. Главным недостатком подхода является увеличение вычислительной сложности, которая, однако, может быть уменьшена на уровне реализации. Алгоритм субполосной декомпозиции в слуховых аппаратах не обязан обладать свойством полного восстановления (*perfect reconstruction*), что позволяет использовать фильтр-прототип с большим ослаблением в полосе заграждения.

Частотный анализ с использованием неравнополосного косинусно-модулированного банка фильтров является основой для построения алгоритмов шумоподавления, компенсации потери слуха, а также компрессии динамического диапазона. Вместе эти алгоритмы образуют блок модификации спектра сигнала (БМСС) – см. рис. 1.10. Общая структура БМСС приведена на рис. 5.1.

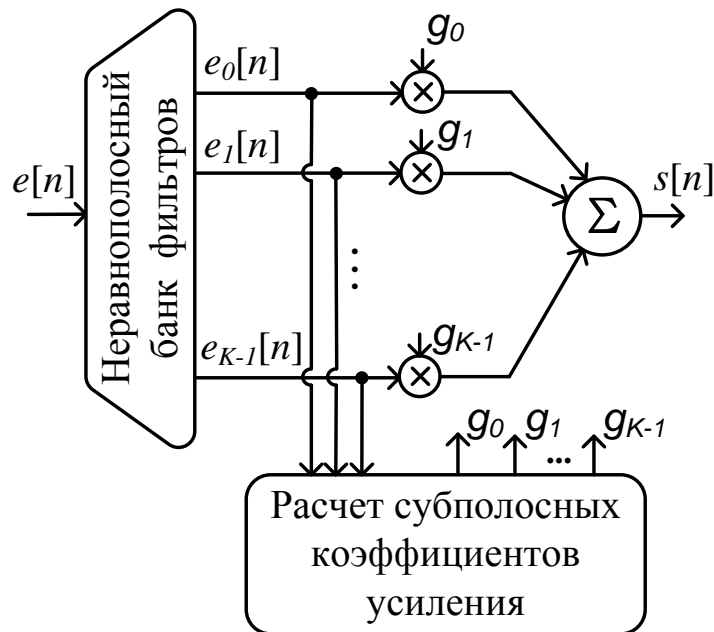


Рис. 5.1. Блок модификации спектра сигнала на основе неравнополосного банка фильтров

5.1.1. Принцип работы

Пусть $G(e^{j\omega})$ представляет собой амплитудную характеристику функции модификации спектра в слуховом аппарате для текущего момента времени. Если $W_n(e^{j\omega})$ частотная характеристика шумоподавителя, то $G(e^{j\omega})$ определяется как [6]

$$G(e^{j\omega}) = B_l(e^{j\omega})W_n(e^{j\omega})C(e^{j\omega}), \quad (5.1)$$

где $C(e^{j\omega})$ – частотная характеристика компрессора динамического диапазона;

$B_l(e^{j\omega})$ – частотная характеристика функции компенсации потери слуха.

Предположим, что $G(e^{j\omega})$ задана на K различных частотах ω_i , $i = 0 \dots K - 1$, которые соответствуют центральным частотам неравнополосного банка фильтров. В этом случае субполосные коэффициенты усиления g_k (рис. 5.1) выбираются таким образом, чтобы минимизировать среднеквадратичную ошибку аппроксимации частотной характеристики слухового аппарата $G(e^{j\omega})$:

$$\epsilon = \sum_{i=-K+1}^{K-1} \left\| G_m(e^{j\omega_i}) - \sum_{k=0}^{K-1} g_k H_k(e^{j\omega_i}) \right\|^2, \quad (5.2)$$

где $H_k(e^{j\omega})$ – частотная характеристика k -го фильтра;

$$\omega_{-i} = -\omega_i;$$

$G_m(e^{j\omega})$ – минимально-фазовая система с такой же АЧХ как у $G(e^{j\omega})$.

Частотную характеристику $G_m(e^{j\omega})$ можно получить путем спектральной факторизации $G(e^{j\omega})$. В данной работе использовался вычислительно эффективный алгоритм [116] на основе дискретного преобразования Гильберта. Известно, что среди всех детерминированных стабильных систем, имеющих одинаковую АЧХ, минимально-фазовая система имеет минимальную групповую задержку [42]. Поэтому использование в качестве целевой минимально-фазовую систему $G_m(e^{j\omega})$ сокращает общую групповую задержку БМСС на рис. 5.1. Решение задачи (5.2) имеет аналитический вид, оно основано на методе «нормальных уравнений» [117]

$$\mathbf{g} = (\mathbf{H}^H \mathbf{H})^{-1} \mathbf{H}^H \mathbf{G}, \quad (5.3)$$

где \mathbf{g} – вектор-столбец, состоящий из субполосных коэффициентов усиления g_0, \dots, g_{K-1} ;

\mathbf{H}^H – матрица эрмитово-сопряженная матрице \mathbf{H} ;

\mathbf{G} и \mathbf{H} – вектор и матрица, определяемые следующим образом:

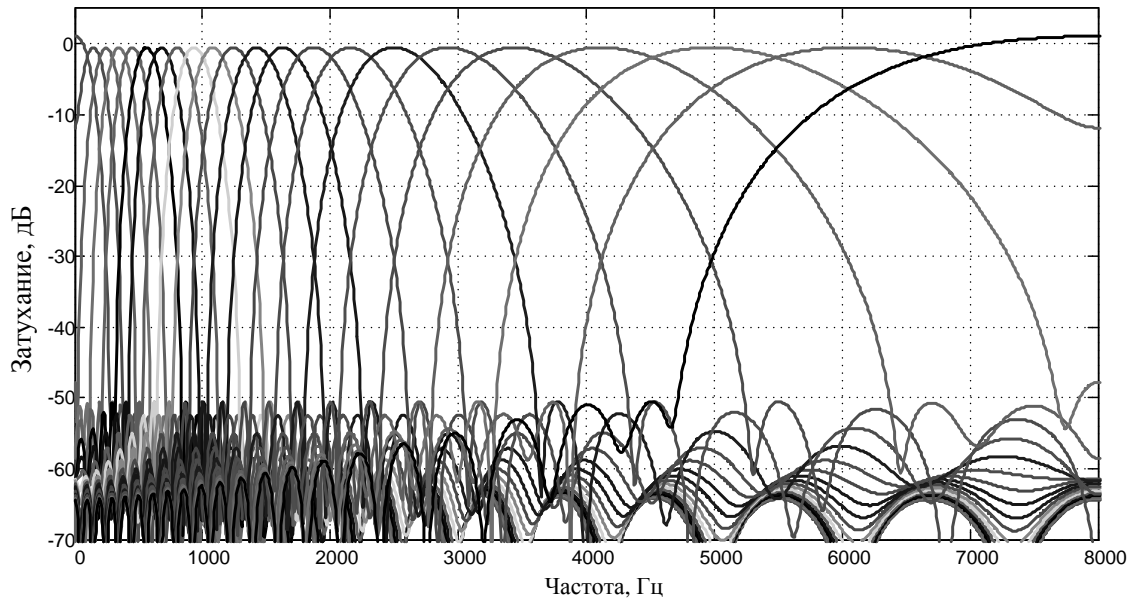


Рис. 5.2. АЧХ 22-канального неравнополосного косинусно-модулированного банка фильтров

$$\mathbf{G} = \left[G_m(e^{-j\omega_{K-1}}) \quad \dots \quad G_m(e^{j\omega_{K-1}}) \right]^T ;$$

$$\mathbf{H} = \begin{bmatrix} H_0(e^{-j\omega_{K-1}}) & H_1(e^{-j\omega_{K-1}}) & \dots & H_{K-1}(e^{-j\omega_{K-1}}) \\ H_0(e^{-j\omega_{K-2}}) & H_1(e^{-j\omega_{K-2}}) & \dots & H_{K-1}(e^{-j\omega_{K-2}}) \\ \vdots & \vdots & \vdots & \vdots \\ H_0(e^{j\omega_0}) & H_1(e^{j\omega_0}) & \dots & H_{K-1}(e^{j\omega_0}) \\ \vdots & \vdots & \vdots & \vdots \\ H_0(e^{j\omega_{K-1}}) & H_1(e^{j\omega_{K-1}}) & \dots & H_{K-1}(e^{j\omega_{K-1}}) \end{bmatrix} .$$

Матрицу $(\mathbf{H}^H \mathbf{H})^{-1} \mathbf{H}^H$ в выражении (5.3) можно рассчитать заранее для набора банка фильтров на заданных частотах ω_i . Эта процедура выполняется один раз на этапе подстройки слухового аппарата. Для выполнения частотного анализа и разбиения сигнала на субполосные составляющие предлагается использовать 22-канальный неравнополосный косинусно-модулированный банк фильтров,¹ АЧХ которого показана на рис. 5.2. Для данного банка фильтров рассчитан фильтр-прототип 44-го порядка с ис-

¹Альтернативным образом, для частотного анализа сигнала в слуховом аппарате может выступать кохлеарный банк фильтров, приведенный в приложении В.

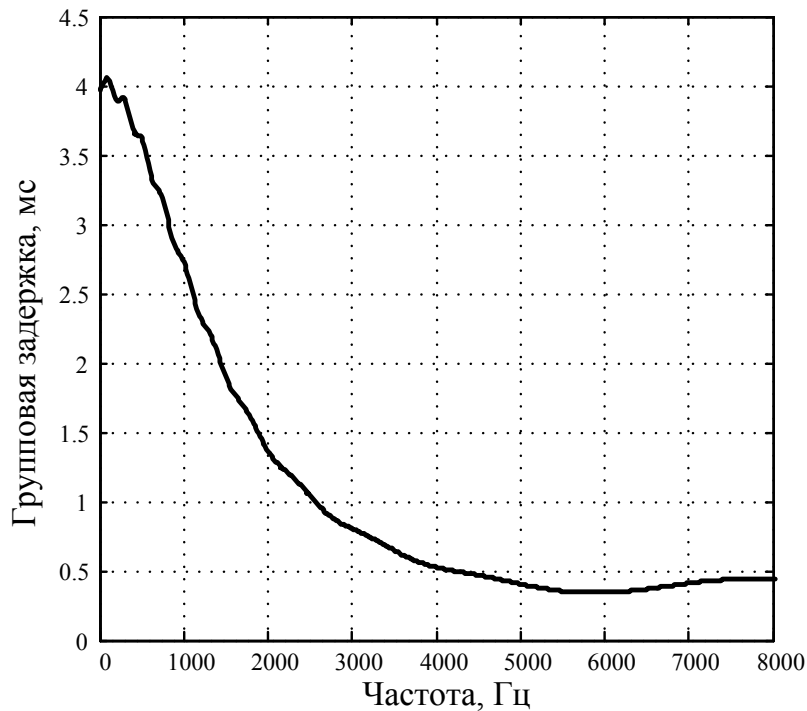


Рис. 5.3. Групповая задержка 22-канального неравнополосного косинусно-модулированного банка фильтров

пользованием метода [75]. Коэффициент деформации частотной оси выбран равным $\alpha = 0,5354$ по методу [47] таким образом, чтобы неравнополосный банк фильтров аппроксимировал психоакустическую шкалу барков. В результате получен банк фильтров, который обладает свойством неравнополосности, имеет малую групповую задержку (рис. 5.3) и достаточное ослабление в полосе заграждения.

5.1.2. Шумоподавление

Функция шумоподавления реализуется с использованием психоакустически-мотивированного правила спектрального взвешивания [118, 119]. Применяемый алгоритм в отличие от многих других, используемых в области телекоммуникаций, имеет целью не полное удаление шума из речевого сигнала, а лишь ослабление шума до приемлемого уровня. Для этой цели в алгоритме имеется настраиваемый параметр $\zeta = 10^{-RL/20}$, который определяет желаемый уровень шума RL в децибелах.

Описание алгоритма. Предполагается, что чистый речевой сигнал и окружающий шум статистически независимы, тогда зашумленный ре-

чевой сигнал $e[n]$ можно представить в виде суммы этих двух компонент. Пусть $S(e^{j\omega_i})$ и $N(e^{j\omega_i})$ это Фурье-образы чистой речи и аддитивного шума, оцененные на частотах ω_i , $i = 0 \dots K - 1$, которые соответствуют центральным частотам неравнополосного банка фильтров. В этом случае желаемый речевой сигнал $\tilde{S}(e^{j\omega_i})$ и шум $\tilde{R}(e^{j\omega_i})$, который будет удален, выражаются как

$$\begin{aligned}\tilde{S}(e^{j\omega_i}) &= S(e^{j\omega_i}) + \zeta N(e^{j\omega_i}); \\ \tilde{N}(e^{j\omega_i}) &= (1 - \zeta)N(e^{j\omega_i}),\end{aligned}$$

а их спектральные плотности мощности (СПМ) имеют вид:

$$\begin{aligned}R_{\tilde{s}}(e^{j\omega_i}) &= R_s(e^{j\omega_i}) + \zeta^2 R_n(e^{j\omega_i}); \\ R_{\tilde{n}}(e^{j\omega_i}) &= (1 - \zeta)^2 R_n(e^{j\omega_i}),\end{aligned}$$

где $R_s(e^{j\omega_i})$ – СПМ речевого сигнала;

$R_n(e^{j\omega_i})$ – СПМ шума.

При использовании алгоритма спектрального взвешивания оценка $\hat{S}(e^{j\omega_i})$ является результатом умножения входного сигнала $S(e^{j\omega_i}) + N(e^{j\omega_i})$ на вещественнозначную взвешивающую функцию $W(e^{j\omega_i})$, имеющую ограничение $0 \leq \zeta \leq W(e^{j\omega_i}) \leq 1$. В этом случае СПМ выходного сигнала будет иметь вид

$$R_{\hat{s}}(e^{j\omega_i}) = W^2(e^{j\omega_i})(R_s(e^{j\omega_i}) + R_n(e^{j\omega_i})).$$

Рассмотрим ошибку $E(e^{j\omega_i})$ между оцененной речью $\hat{S}(e^{j\omega_i})$ и желаемым выходным сигналом $\tilde{S}(e^{j\omega_i})$. СПМ ошибки можно выразить в виде суммы двух компонент $R_E(e^{j\omega_i}) = R_{E_s}(e^{j\omega_i}) + R_{E_n}(e^{j\omega_i})$, где

$$\begin{aligned}R_{E_s}(e^{j\omega_i}) &= R_s(e^{j\omega_i})(W(e^{j\omega_i}) - 1)^2, \\ R_{E_n}(e^{j\omega_i}) &= R_n(e^{j\omega_i})(W(e^{j\omega_i}) - \zeta)^2.\end{aligned}$$

Компонента $R_{E_s}(e^{j\omega_i})$ описывает искажение речи, которое минимизируется путем выбора взвешивающего множителя $W(e^{j\omega_i})$ равным 1. Вторую компоненту $R_{E_n}(e^{j\omega_i})$ можно интерпретировать как разницу между желаемой мощностью шума и фактической мощностью. Эту компоненту можно минимизировать, устанавливая множитель $W(e^{j\omega_i}) = \zeta$. Поскольку обе компоненты квадратично зависят от $W(e^{j\omega_i})$, то минимальное значение $R_E(e^{j\omega_i})$

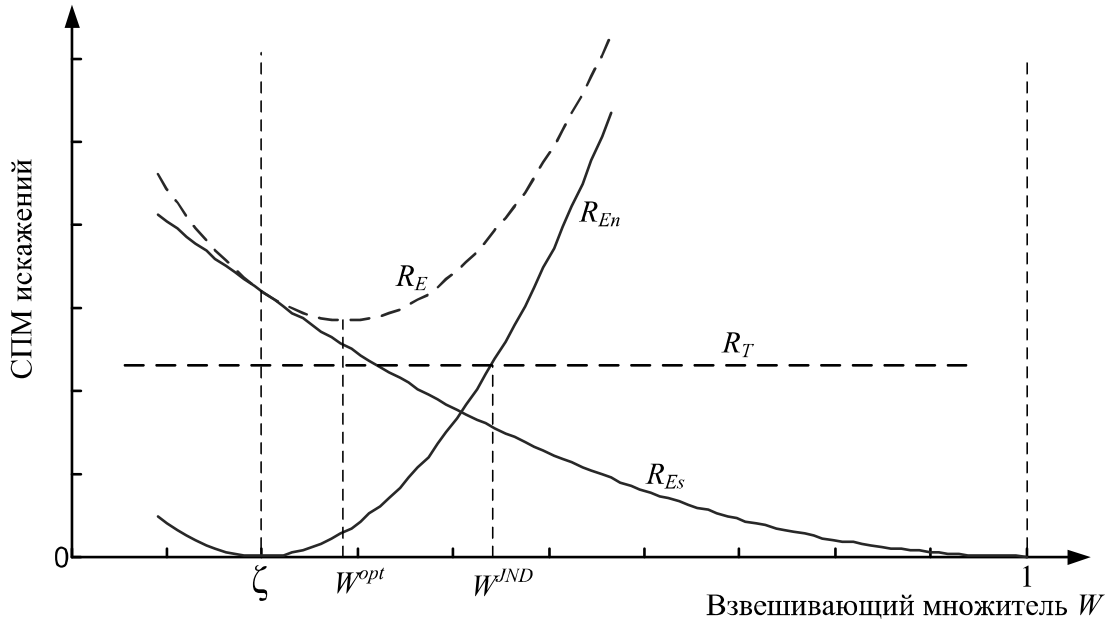


Рис. 5.4. Компоненты искажений речи, появляющиеся в результате спектрального взвешивания

будет достигаться при некотором $W^{\text{opt}}(e^{j\omega_i})$ (рис. 5.4). Фактически при $\zeta = 0$ решение $W^{\text{opt}}(e^{j\omega_i})$ соответствует винеровскому фильтру [49].

В [118] для получения перцептуально-мотивированного правила спектрального взвешивания предлагается выбирать $W(e^{j\omega_i})$ таким образом, чтобы замаскировать остаточный шум, допуская при этом незначительное искажение речи. Для этого достаточно сохранять $R_{E_n}(e^{j\omega_i})$ равным порогу маскирования. Это решение достигается, если выбрать $W(e^{j\omega_i})$ таким, что СПМ ошибки ($R_E(e^{j\omega_i})$) в точности подпадает под порог маскирования:

$$R_{E_n}(e^{j\omega_i}) = R_n(e^{j\omega_i})(W(e^{j\omega_i}) - \zeta)^2 = R_T(e^{j\omega_i}).$$

Решая это уравнение относительно $W(e^{j\omega})$ при условии $\zeta \leq W(e^{j\omega_i}) \leq 1$, получаем искомое правило взвешивания²

$$W^{\text{JND}}(e^{j\omega_i}) = \min\left(\sqrt{\frac{R_T(e^{j\omega_i})}{R_n(e^{j\omega_i})}} + \zeta, 1\right). \quad (5.4)$$

Найденная таким образом функция $W^{\text{JND}}(e^{j\omega})$ соответствует функции шумоподавления $W_n(e^{j\omega})$ в выражении (5.1).

²JND (*Just Notable Distortion*) – едва заметное искажение.

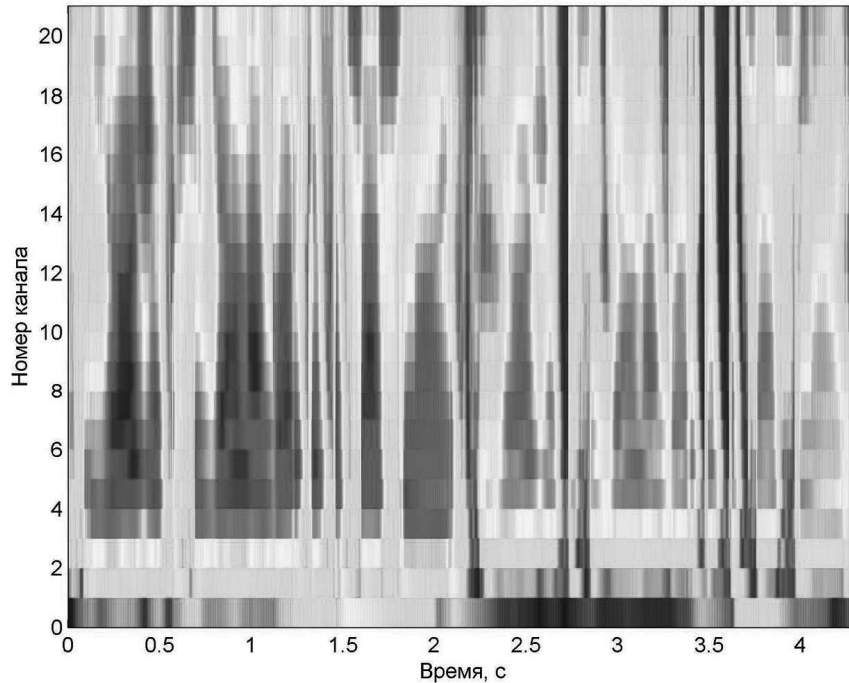


Рис. 5.5. Оценка СПМ при помощи неравнополосного банка фильтров (постоянная времени усреднения $T = 4$ мс)

Оценка СПМ зашумленного сигнала $R_e(e^{j\omega_i})$ в момент времени n получается при помощи экспоненциального усреднения [120] выходных сигналов неравнополосного банка фильтров:

$$R_e^{(n)}(e^{j\omega_k}) = R_e(n, k) = \alpha R_e(n-1, k) + (1 - \alpha) e_k^2[n]; \quad \alpha = e^{-\frac{1}{f_s T}},$$

где $e_k[n]$ – выход k -го фильтра в момент времени n ;

T – требуемое постоянная времени усреднения;

f_s – частота дискретизации.

Пример оценки СПМ речевого сигнала при помощи 22-канального банка неравнополосного банка фильтров показан на рис. 5.5.

Для получения СПМ шума $R_n(e^{j\omega_i})$, которая необходима для расчета взвешивающей функции (5.4), используется эффективный робастный алгоритм оценки шума с экспоненциальным усреднением и контролем по минимуму энергии [121] (*MCRA – Minima Controlled Recursive Averaging*). Этот подход занимает промежуточное положение между обычным экспоненциальным усреднением и методом минимальной статистики [122]. СПМ оценивается в каждом канале k путем экспоненциального усреднения зна-

чений $R_e(n, k)$ и использует постоянную усреднения, которая регулируется в зависимости от вероятности наличия речи. Во время работы слухового аппарата взвешивающая функция (5.4) пересчитывается каждые 4 мс.

Оценка производительности системы шумоподавления. Для оценки производительности системы шумоподавления используется четыре записи с мужским и женским голосами общей длительности 25 с. В эксперименте к сигналам добавлялся белый шум в полосе частот 500–4000 Гц. Производительность оценивалась при пяти различных отношениях сигнал/шум (ОСШ). В качестве критерия качества работы шумоподавителя выбран объективный параметр SegSNR (*segmented signal-to-noise ratio* – сегментированное ОСШ), который определяется как среднее значение ОСШ каждого выходного фрейма сигнала системы шумоподавления слухового аппарата:

$$\text{SegSNR} = \frac{1}{B} \sum_{m=1}^B 10 \log_{10} \left(\frac{P_{\text{signal}}^{(m)}}{P_{\text{noise}}^{(m)}} \right),$$

где B – число фреймов сигнала;

$P_{\text{signal}}^{(m)}$ – мощность сигнала на m -м фрейме;

$P_{\text{noise}}^{(m)}$ – мощность шума на m -м фрейме.

В экспериментах размер фрейма выбирался равным 30 мс. Полезный сигнал получен как выход слухового аппарата при пропуске чистого речевого сигнала. Шум оценивался как разность между полезным сигналом и выходом слухового аппарата с работающей системой шумоподавления, при подаче на вход зашумленного речевого сигнала.

В качестве второго критерия качества выбран индекс разборчивости речи. Измерение разборчивости производилось по методу SII (*speech intelligibility index*), который основан на вычислении индекса артикуляции. В соответствии с методикой [123] весь частотный диапазон речевого сигнала разделяется на 18 полос, в пределах которых определяется ОСШ. Сигнал делится на полосы такой ширины, чтобы вклад каждой полосы в результирующую разборчивость был одинаковый (полосы одинаковой разборчивости). В каждой полосе ОСШ рассчитывается, определенным обра-

Таблица 5.1. Производительность системы шумоподавления слухового аппарата

	Входное отношение сигнал/шум				
	20 дБ	15 дБ	10 дБ	5 дБ	0 дБ
SII (до обработки)	0,73	0,62	0,52	0,47	0,37
SII (после обработки)	0,74	0,66	0,59	0,51	0,41
SegSNR, дБ	28,5	23,1	19,0	14,1	7,9

зом взвешивается и используется для вычисления коэффициента SII, который может изменяться в пределах от 0 до 1. Значения индекса принято интерпретировать следующим образом [123]:

- ниже 0,3 – плохая разборчивость;
- от 0,3 до 0,5 – удовлетворительная разборчивость;
- от 0,5 до 0,7 – хорошая разборчивость;
- выше 0,7 – очень хорошая разборчивость.

В табл. 5.1 приведены результаты экспериментов: усредненные значения разборчивости и получаемые ОСШ. Приведенные результаты показывают, что система шумоподавления позволяет получить значительное улучшение ОСШ, повышая при этом индекс разборчивости речи. На рис. 5.6 показан пример работы шумоподавителя.

5.1.3. Компенсация потери слуха

Порог слышимости – это минимальная величина звукового давления, при которой звук данной частоты может быть воспринят ухом человека. При нарушении слуха порог слышимости может значительно изменяться. Порог слышимости зависит от частоты звука. При действии шумов и других звуковых раздражителей порог слышимости для данного звука повышается (маскировка звука), причём повышенное значение порога слышимости сохраняется некоторое время после прекращения действия мешающего фактора, а затем постепенно возвращается к исходному уровню. У разных людей и у одних и тех же лиц в разное время порог слышимости может

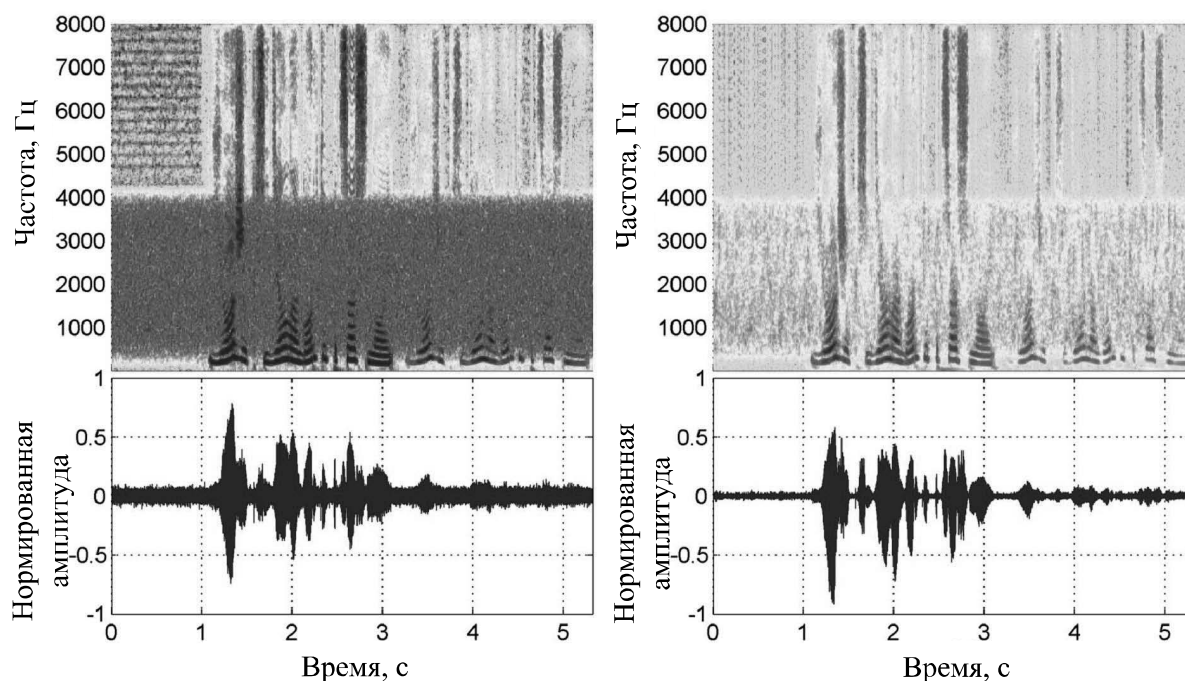


Рис. 5.6. Пример работы системы шумоподавления слухового аппарата (параметр $RL = 15$)

различаться в зависимости от возраста, физиологического состояния, тренированности.

Определение порога слышимости для компенсации потери слуха, как правило производят при помощи тональных аудиограмм. Аудиограмма – это график, отображающий состояние слуха человека. По горизонтальной оси откладываются частоты (набор значений частот стандартный – 250, 500, 1000, 2000, 3000, 4000 и 8000 Гц), а по вертикальной – пороги слышимости на соответствующих частотах, т.е. минимальные уровни звукового давления сигнала, при которых пациент слышит звук. При построении аудиограммы значения этих порогов измеряются специальным прибором – аудиометром [2, 12].

По имеющейся аудиограмме для каждой частоты рассчитывается коэффициенты компенсации потери слуха (*insertion gain*). Форма коэффициентов компенсации потери слуха зависит от метода корректировки слуха (*prescription method*) и не обязательно полностью повторяет форму аудиограммы. На рис. 5.7 показаны пример типичной аудиограммы и рассчитанные коэффициенты компенсации потери слуха. Также на рисунке показано, как предлагаемый блок модификации спектра сигнала на основе неравно-

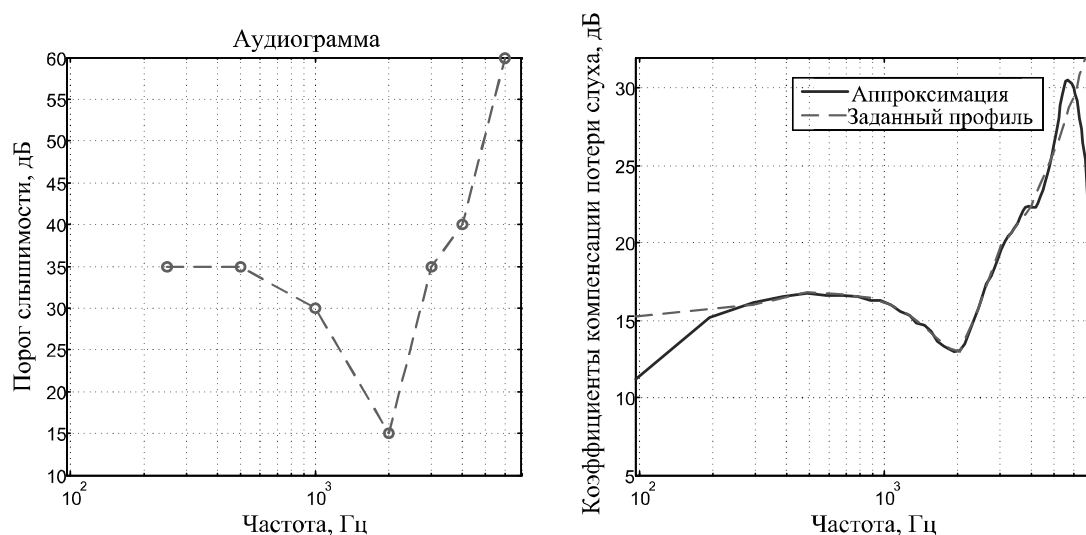


Рис. 5.7. Пример аппроксимации коэффициентов компенсации потери слуха при помощи блока модификации спектра сигнала (ось частот в логарифмическом масштабе)

полосного банка фильтров позволяет аппроксимировать заданную форму коэффициентов компенсации потери слуха.

5.1.4. Компрессия динамического диапазона

Компрессия динамического диапазона (КДД) используется для уменьшения разницы в уровнях компонент с большой и малой интенсивностью в аудиосигнале. В слуховых аппаратах КДД применяется для отображения широкого динамического диапазона речевого сигнала в суженный динамический диапазон остаточного слуха [124]. Люди с потерей слуха, как правило, страдают от понижения порога слышимости, потери разборчивости и уменьшения динамического диапазона. Поэтому для компенсации потери слуха необходимо решить задачу частотно-зависимого усиления сигнала и компрессию динамического диапазона. Специфическим требованием для слуховых аппаратов является малая алгоритмическая задержка на обработку сигнала. Для удовлетворения указанным требованиям для частотного анализа сигнала используется неравнополосный банк фильтров.

Принцип динамической компрессии. Главной задачей компрессора является автоматический выбор коэффициента усиления субполосных

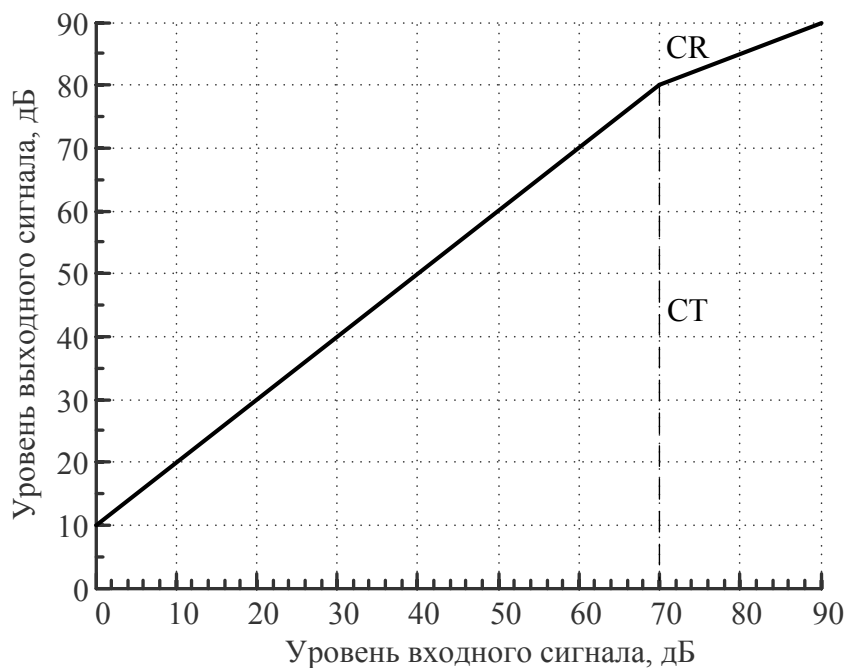


Рис. 5.8. Характеристика вход/выход компрессора

сигналов в зависимости от уровня мощности сигнала в данной полосе. Субполосы с высокой мощностью ослабляются, а с малой мощностью усиливаются. Благодаря такой обработке тихие звуки становятся слышимыми, а громкие звуки не вызывают дискомфортных ощущений. КДД, как правило, имеет следующие параметры [124]:

- порог компрессии (СТ);
- коэффициент сжатия (CR);
- время атаки и отпускания;
- коэффициент усиления слухового аппарата (G_{dB}).

Порог компрессии (СТ), измеряемый в децибелах, определяет точку изгиба характеристики вход/выход компрессора, после которой алгоритм КДД становится активным. Если уровень входного сигнала ниже СТ, то выходной сигнал усиливается линейно. В случае, когда уровень выходного сигнала выше порога компрессии (СТ), коэффициент усиления компрессора уменьшается. Параметр CR определяет степень сжатия динамического диапазона. Например, значение $CR=5$ (или 5:1) означает, что на каждые 5 дБ увеличения уровня входного сигнала уровень выходной сигнал возрастет на 1 дБ. На рис. 5.8 приводится характеристика вход/выход компрессора для параметров $CR = 2$, $СТ = 70$ дБ и $G_{dB} = 10$ дБ.

Алгоритм КДД выполняется отдельно в каждом канале неравнополосного банка фильтров. Результатом работы алгоритма является вещественнозначная функция $C(e^{j\omega_k})$, которая показывает, какой коэффициент усиления должен быть применен в k -м канале банка фильтров (ω_k – центральная частота k -ого фильтра).

Алгоритм КДД выполняет оценку мощности сигнала в каждой полосе при помощи экспоненциального усреднения [120]. Но в отличие от алгоритма шумоподавления мощность оценивается в двух режимах – атаки и отпускания. Для этого рассчитываются две константы усреднения α и β , зависящие от времени атаки и времени отпускания. Таким образом, мощность сигнала в алгоритме КДД вычисляется следующим образом:

$$P(k, n) = \begin{cases} \alpha P(k, n-1) + (1-\alpha)e_k^2[n], & \text{если } P(k, n-1) < e_k^2[n]; \\ \beta P(k, n-1) + (1-\beta)e_k^2[n], & \text{иначе.} \end{cases}$$

где $P(k, n)$ – мощность сигнала в k -м канале неравнополосного банка фильтров в момент времени n ;

$e_k[n]$ – выход k -го фильтра в момент времени n .

Типичное значение времени атаки и отпускания составляет 10 и 100 мс соответственно. Это позволяет компрессору быстро реагировать на увеличение мощности, но тем не менее не компрессировать короткие пики. Когда уровень сигнала падает в компрессоре на время отпускания, сохраняется высокий уровень на случай, если это короткая пауза между громкими звуками.

Оцененные в субполосах уровни мощности переводятся в децибелы $P_{k,\text{dB}}^{\text{in}} = 10 \log_{10}(P(k, n))$ и выполняется расчет *ожидаемых* уровней мощности с использованием характеристики вход/выход компрессора [125]:

$$P_{k,\text{dB}}^{\text{out}} = \begin{cases} P_{k,\text{dB}}^{\text{in}} + G_{\text{dB}}, & \text{если } P_{k,\text{dB}}^{\text{in}} < \text{CT}, \\ \text{CT} + \frac{1}{\text{CT}}(P_{k,\text{dB}}^{\text{in}} - \text{CT}) + G_{\text{dB}}, & \text{иначе.} \end{cases}$$

Далее рассчитываются коэффициенты усиления для каждого канала ($k = 0, \dots, K-1$) неравнополосного банка фильтров:

$$C(e^{j\omega_k}) = 10^{(P_{k,\text{dB}}^{\text{out}} - P_{k,\text{dB}}^{\text{in}})/10}.$$

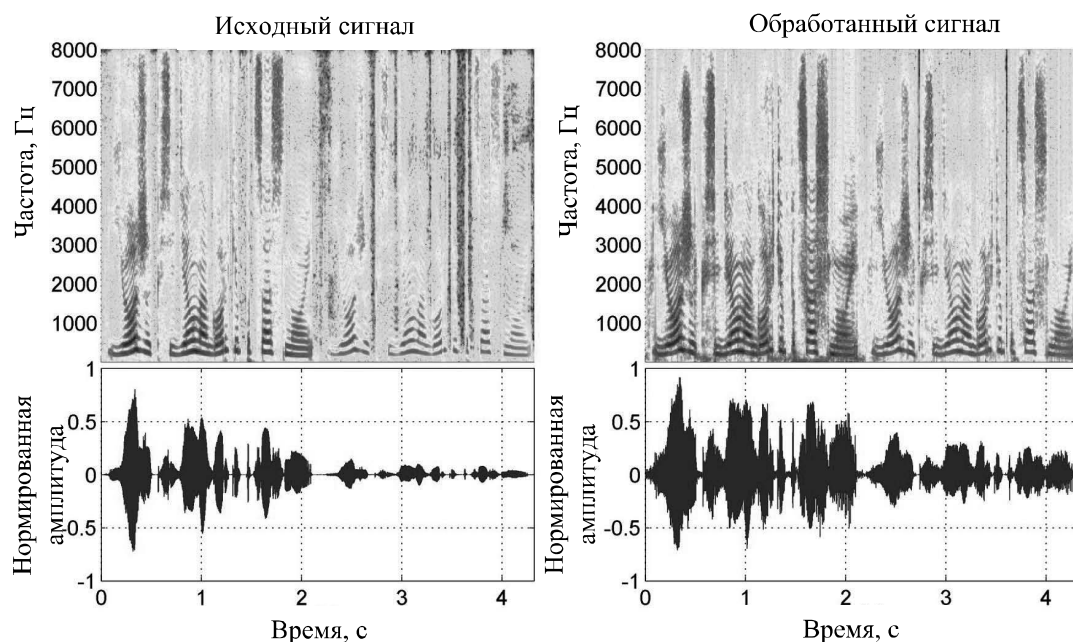


Рис. 5.9. Пример обработки сигнала при помощи алгоритма КДД

Проверка предлагаемого алгоритма многополосной компрессии динамического диапазона производилась на речевой базе, состоящей из мужских и женских голосов. На рис. 5.9 показан пример одного сигнала, состоящего из двух частей – громкого и тихого участков. После обработки в компрессоре обе части приблизительно выровнялись по громкости. Искажения, видимые в спектре после обработки, появляются в результате нелинейной обработки в компрессоре, не оказывают ощутимого влияния на разборчивость речи и узнаваемость диктора.

На рис. 5.10 представлена спектральная гистограмма тестового аудиофайла, содержащего чистую речь. Каждый вертикальный срез представляет собой гистограмму значений уровня сигнала (в дБ) в соответствующей частотной полосе. Темный цвет говорит о более высокой плотности вероятности. На рисунке также показаны границы динамического диапазона для каждой частотной полосы, измеренные для человека с нормальным слухом. Тестовый речевой сигнал обрабатывался при помощи предлагаемого алгоритма многополосной компрессии динамического диапазона, настроенного на конкретный профиль потери слуха. Получаемая после обработки сигнала спектральная гистограмма представлена на рис. 5.11. Результаты показывают, что предлагаемый алгоритм позволяет адаптировать уровень выходного сигнала к слуховой характеристике пациента.

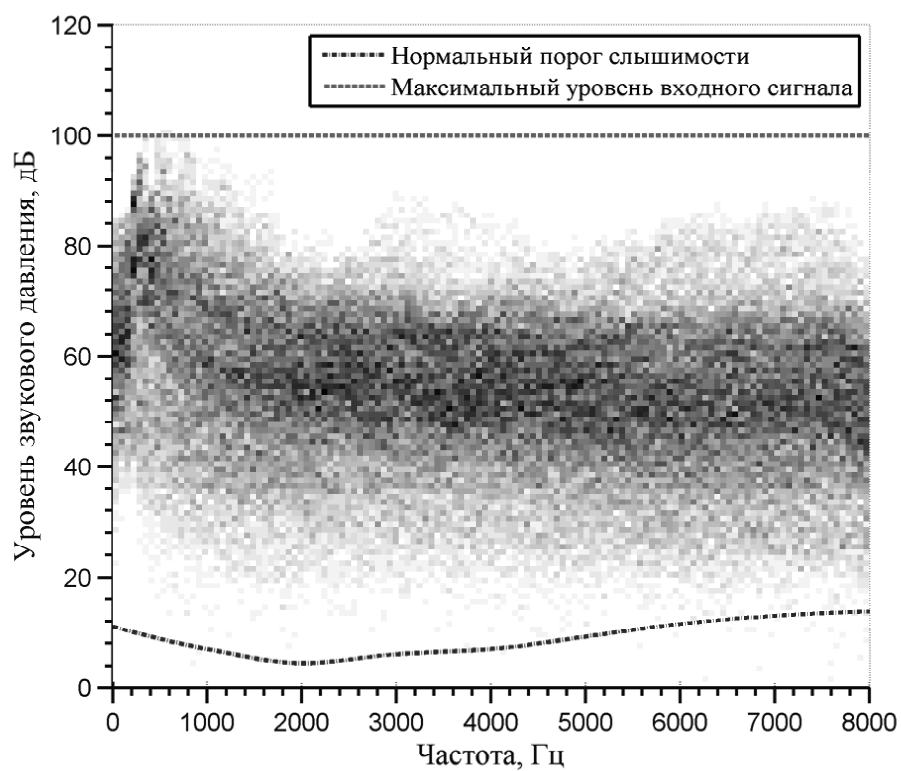


Рис. 5.10. Спектральная гистограмма речевого сигнала

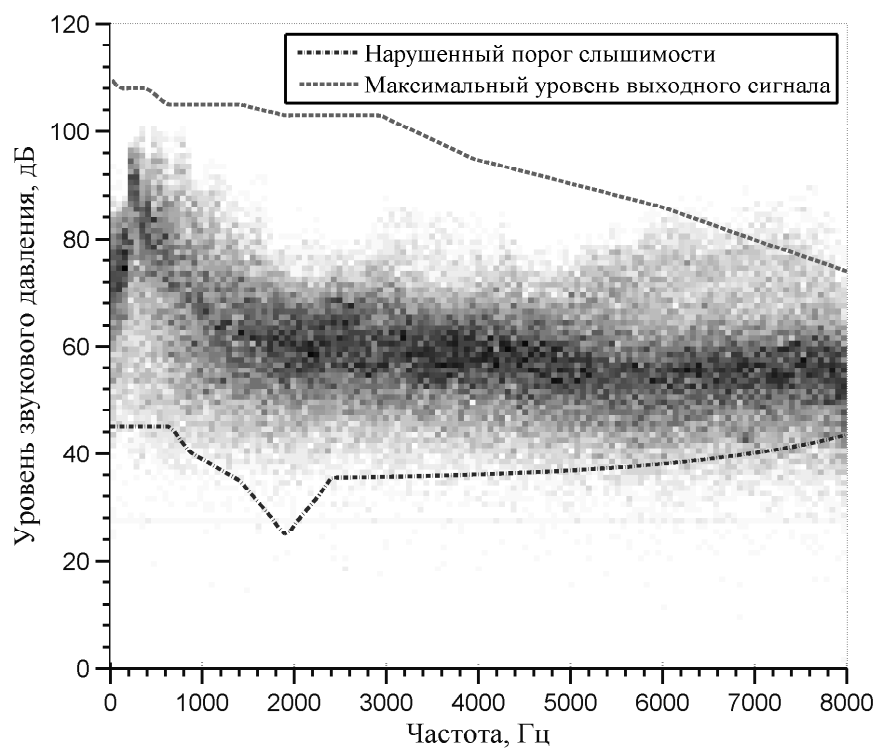


Рис. 5.11. Спектральная гистограмма сигнала, обработанного при помощи алгоритма КДД

5.2. Компенсация эффекта акустической обратной связи с использованием неравнополосного банка фильтров

5.2.1. Введение

Утечка сигнала из динамика слухового аппарата в микрофон, называемая акустической обратной связью (АОС), значительно ограничивает максимальный коэффициент усиления, который может быть использован в слуховом аппарате при условии сохранения его устойчивости. АОС вызывает эффект завывания и значительно ухудшает качество звука, когда коэффициент усиления близок к максимальному. В ряде случаев максимального усиления недостаточно, чтобы компенсировать потерю слуха. Для уменьшения влияния АОС предложено несколько методов, использующих адаптивные алгоритмы. Наибольшее распространение в практических реализациях слуховых аппаратов получили адаптивные алгоритмы на основе метода наименьших средних квадратов благодаря их простоте и малой вычислительной сложности [26]. Тем не менее эти алгоритмы не всегда позволяют обеспечить достаточную устойчивость для широкополосного сигнала. Эффективным способом устранения данного недостатка является использование в системе компенсации эффекта АОС банка фильтров. При помощи банка фильтров сигнал может быть расщеплен на субполосные составляющие, и процесс адаптации может выполняться отдельно в каждой субполосе. За счет этого достигается большая устойчивость системы, а также увеличивается скорость сходимости адаптивного алгоритма. Дополнительными преимуществами такого подхода являются возможность уменьшения порядка адаптивных фильтров и сокращение вычислительной сложности за счет децимации субполосных сигналов [126].

В [127] предлагается выполнять субполосную декомпозицию сигнала с помощью банка фильтров WOLA для улучшения производительности системы компенсации эффекта АОС (рис. 5.12). На схеме индексы n и m используются для обозначения временных отсчетов исходного (широкополосного) сигнала и отсчетов субполосных сигналов соответственно. Жирные линии на схеме используются для обозначения векторных вели-

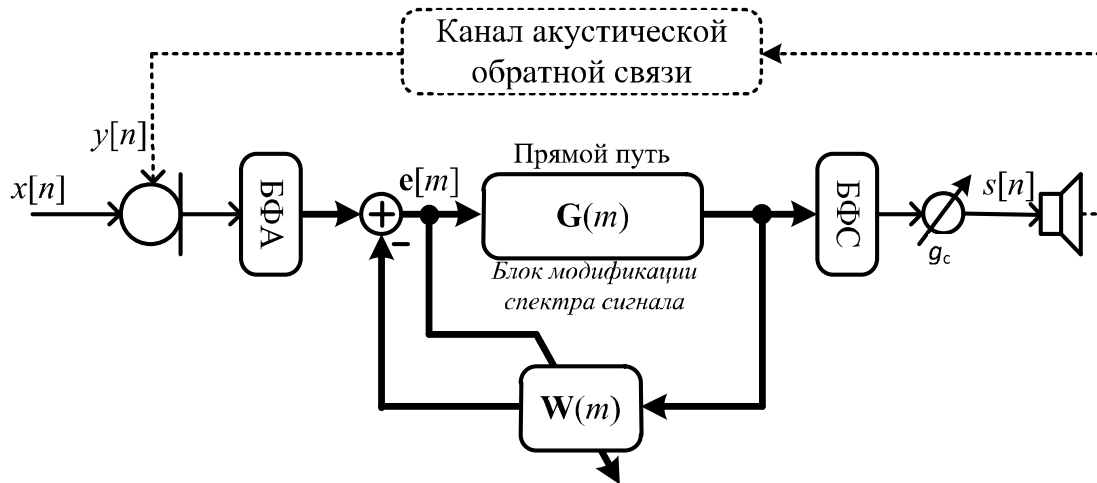


Рис. 5.12. Схема компенсации эффекта АОС с использованием банка фильтров (БФА – банк фильтров анализа, БФС – банк фильтров синтеза)

чин. Блок модификации спектра сигнала (прямой путь слухового аппарата) выполняет функции компенсации потери слуха, подавления шума и компрессии динамического диапазона. Коэффициент g_c применяется для общего усиления выходного уровня звука, он может изменяться пользователем слухового аппарата посредством регулятора громкости.

Недостатком данного подхода является то, что банк фильтров вносит ощутимую задержку (порядка 16 мс) в прямой путь распространения сигнала между микрофоном и динамиком. Это нежелательно по нескольким причинам. Во-первых, при задержке более 10 мс возникает эффект гребенчатого фильтра (*comb filter effect*) [5], который воспринимается человеческим ухом как эхо и является результатом того, что звук, обработанный слуховым аппаратом, накладывается в среднем ухе на необработанный звук. Во-вторых, собственный голос говорящего достигает улитки внутреннего уха посредством костной звукопроводности и через отверстие в слуховом аппарате с минимальной задержкой, где взаимодействует с задержанным и усиленным сигналом от слухового аппарата, что вызывает эффект окрашивания звука (*coloration effect*) [6].

Отсутствие дополнительной задержки в прямом пути можно обеспечить за счет параллельного включения системы компенсации эффекта АОС (рис. 5.13). Однако данная система не позволяет выполнять адаптивную фильтрацию в субполосах сигнала $s[n]$, поскольку в этом случае в

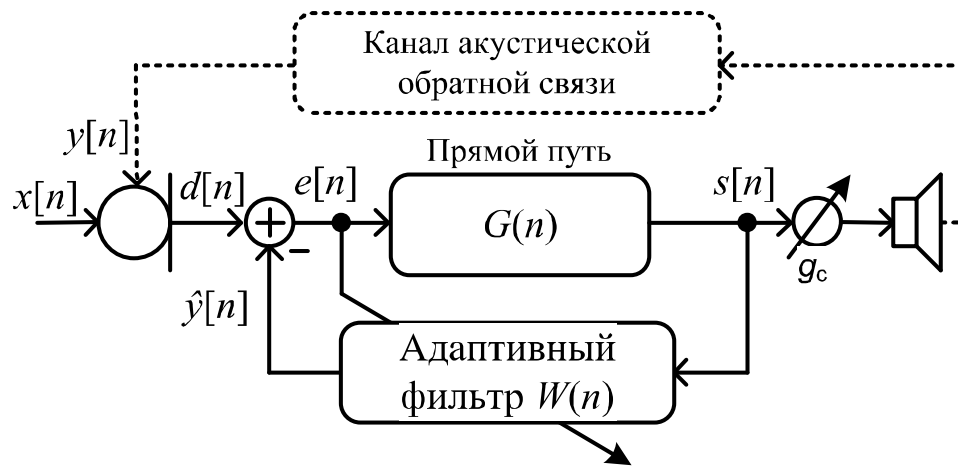


Рис. 5.13. Схема компенсации эффекта АОС, не вносящая задержки в прямой путь распространения сигнала

сигнал компенсации $\hat{y}[n]$ будет внесена задержка, обусловленная банками анализа/синтеза.

5.2.2. Метод субполосной компенсации эффекта акустической обратной связи

В данном разделе разрабатывается метод субполосной компенсации эффекта АОС, который не вносит дополнительной задержки в прямой путь распространения сигнала в слуховом аппарате [51]. Работа опирается на последние результаты в области адаптивной фильтрации, которые показывают, что неравнополосные адаптивные структуры в ряде случаев превосходят равнополосные по таким параметрам, как скорость сходимости и/или модельная ошибка [50]. В отличие от [127] для субполосной декомпозиции сигнала в работе используется передискретизированный неравнополосный косинусно-модулированный банк фильтров (КМБФ) [75].

Устранение задержки банков фильтров анализа/синтеза. Основная идея предлагаемого подхода заключается в том, чтобы устранить задержку, вносимую в компенсирующий сигнал $\hat{y}[n]$ банками фильтров, за счет добавления в адаптивный фильтр функции предсказания. Данный подход оправдан, поскольку эффект АОС приводит к появлению периодиче-

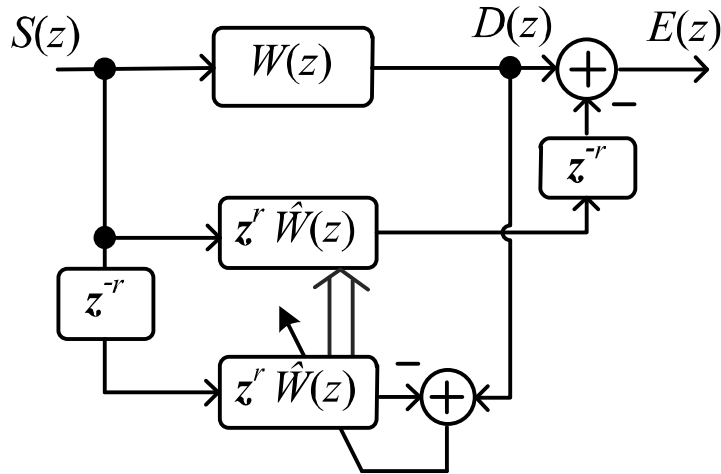


Рис. 5.14. Устранение задержки при помощи адаптивной фильтрации с предсказанием (широкой стрелкой показана операция копирования коэффициентов адаптивного фильтра)

ской помехи на определенной частоте, а, как известно, детерминированные (периодические) сигналы предсказываются с высокой точностью.

Рассмотрим алгоритм адаптивной фильтрации на основе метода наименьших средних квадратов. Пусть $S(z)$ – это z -образ входного сигнала, а $D(z)$ – обучающего сигнала, появляющегося на выходе линейной системы с передаточной функцией $W(z)$, тогда в результате работы адаптивного алгоритма находится оценка передаточной функции $\hat{W}(z)$. Если задержать $S(z)$ на r отсчетов, то система будет выполнять функцию линейного предсказания на r отсчетов, а передаточная функция примет вид $z^r \hat{W}(z)$. Допустим, что банк фильтров синтеза вносит задержку в компенсирующий сигнал, равную r отсчетам, тогда входной и обучающий сигналы можно синхронизировать, используя фильтр с передаточной функцией $z^r \hat{W}(z)$ (рис. 5.14).

Используя изложенный выше принцип устранения задержки, предложим следующую схему субполосной компенсации эффекта АОС [51] (рис. 5.15).

При помощи банков фильтров анализа выполняется расщепление сигналов $d[n]$ и $s[n]$ на M спектральных компонент. Поскольку спектры сигналов внутри каналов занимают более узкие полосы частот, выполняется переход на более низкую частоту дискретизации. В банке фильтров синтеза исходная частота дискретизации восстанавливается. В каждом канале бан-

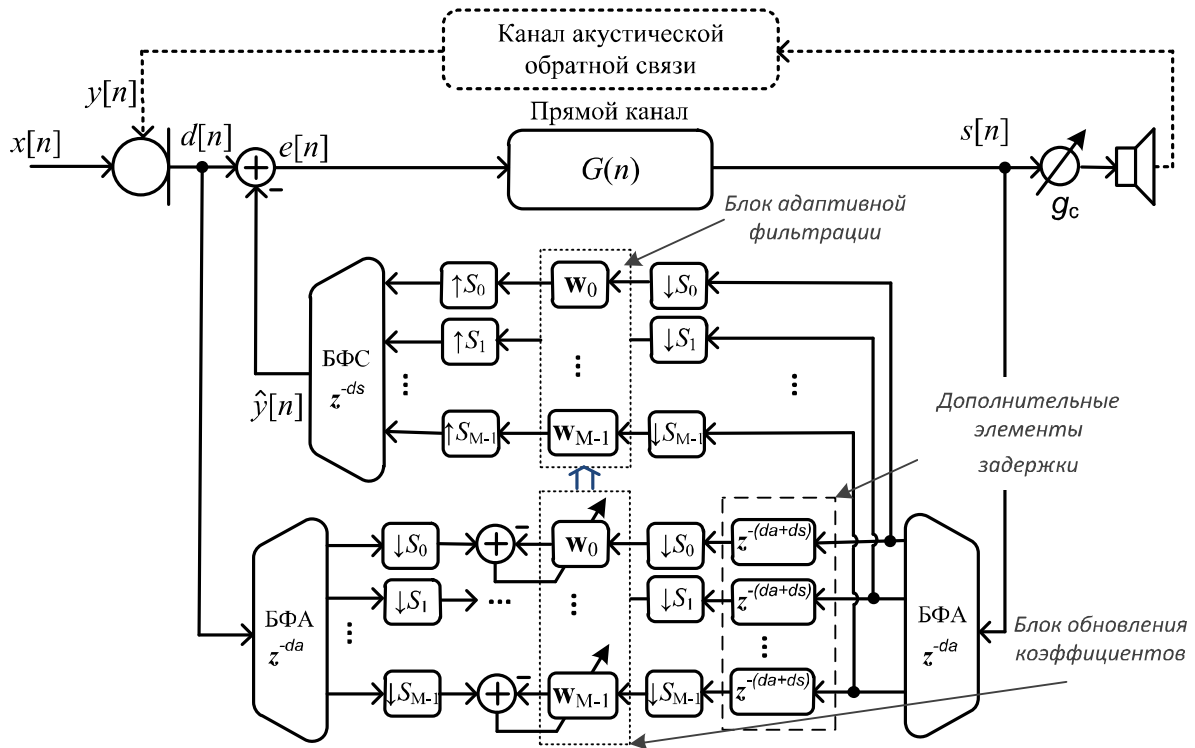


Рис. 5.15. Схема субполосной компенсации акустической обратной связи на основе неравнополосного КМБФ

ка фильтров оценивается свой вектор коэффициентов адаптивного фильтра (на схеме соответствующие векторы обозначены $\mathbf{w}_0, \dots, \mathbf{w}_{M-1}$). Банки фильтров вносят следующие задержки: da – задержка в банке фильтров анализа; ds – задержка в банке фильтров синтеза. С помощью изложенного выше принципа для синхронизации желаемого сигнала $d[n]$ с входным $s[n]$ применяются дополнительные элементы задержки.

Адаптивная фильтрация субполосных сигналов. Особенность алгоритма адаптивной фильтрации, используемого в системе компенсации эффекта АОС, заключается в том, что обновление коэффициентов фильтров и сама фильтрация происходят отдельно: в блоке обновления коэффициентов и блоке адаптивной фильтрации. В каждом канале оценивается свой набор коэффициентов адаптивного фильтра. Процедура оценки одинакова для всех каналов и отличается лишь значениями параметров, таких как порядок фильтра, коэффициент потерь и шаг адаптации. В блоке обновления коэффициентов используется стандартный алгоритм наименьших

квадратов [9] (для упрощения записи индекс номера канала опущен):

1. Каждому коэффициенту фильтра $w[l]$, $l = 0, 1, \dots, L - 1$, присваивается нулевое значение, L – порядок адаптивного фильтра.

2. Вычисляется выходной отсчет фильтра:

$$\hat{y}[m] = \sum_{l=0}^{L-1} w[l]s[m-l],$$

где m – номер текущего отсчета входного сигнала;

$s[m]$ – входной сигнал блока обновления коэффициентов.

3. Вычисляется оценка ошибки:

$$e[m] = d[m] - \hat{y}[m],$$

где $d[m]$ – желаемый сигнал.

4. Обновляются весовые коэффициенты:

$$w[l] = \delta w[l] + 2\mu e[m]s[m-l],$$

где $0 < \delta < 1$ – коэффициент потерь;

μ – шаг адаптации.

Увеличивается номер текущего отсчета: $m = m + 1$. Алгоритм переходит к шагу 2.

В блоке адаптивной фильтрации выполняется только второй шаг алгоритма, а коэффициенты $w[l]$ копируются из блока обновления коэффициентов.

На рис. 5.16 показаны сигнал на входе слухового аппарата и выходные сигналы широкополосной и субполосной систем компенсации эффекта АОС. Из приведенных графиков можно сделать вывод, что субполосная система имеет более высокую скорость сходимости и нежелательные осцилляции в выходном сигнале значительно меньше, чем в случае использования широкополосной системы компенсации эффекта АОС.

В системе компенсации эффекта АОС используется неравнополосный КМБФ, ширина полос которого адаптирована к разрешающей способности слухового восприятия [47]. Исходя из того, что слуховая система человека более чувствительна к изменениям сигнала в области нижних частот,

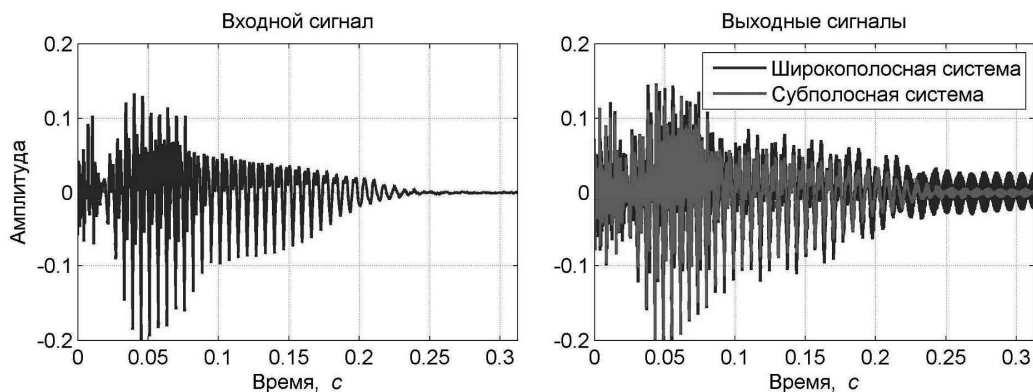


Рис. 5.16. Качественное сравнение широкополосной и субполосной систем подавления эффекта АОС

банк фильтров должен обладать следующим свойством: ширина его полос пропускания должна возрастать с увеличением частоты.

Для использования в субполосной системе компенсации эффекта АОС (см. рис. 5.15) был рассчитан 8-канальный неравнополосный КМБФ по методу [75]. На рис. 5.17–5.18 изображены амплитудно-частотные характеристики полученного банка фильтров, а его групповая задержка – на рис. 5.19. Коэффициенты децимации/интерполяции неравнополосного КМБФ имеют значения $S_k = \{19, 9, 6, 4, 2, 1, 1, 1\}$.

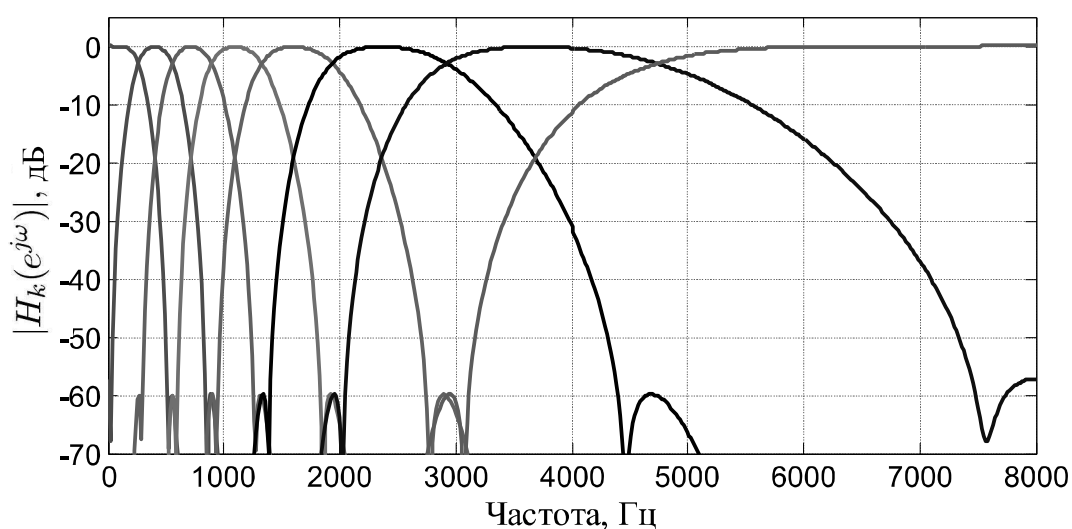


Рис. 5.17. АЧХ банка фильтров анализа

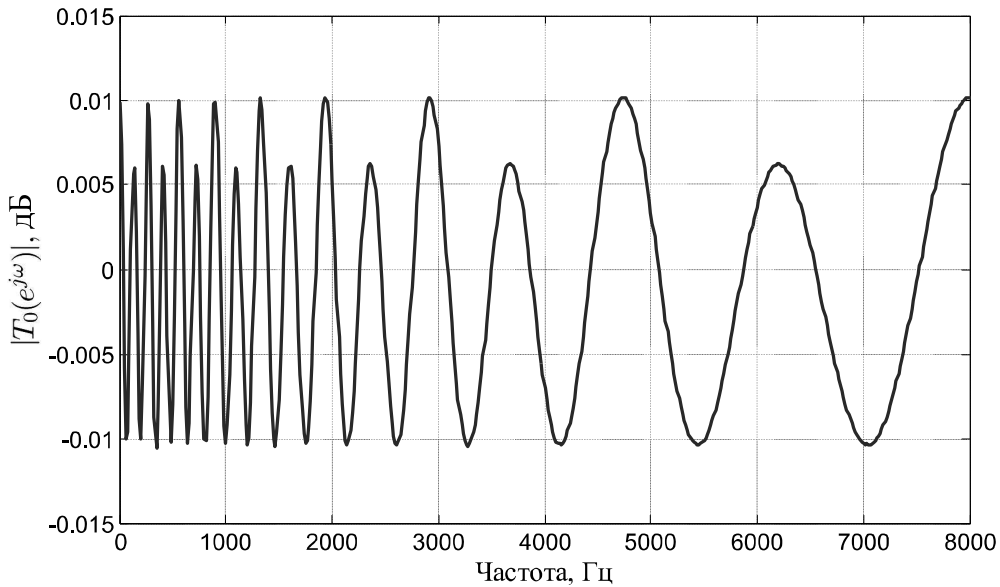


Рис. 5.18. АЧХ системы анализа/синтеза банка фильтров

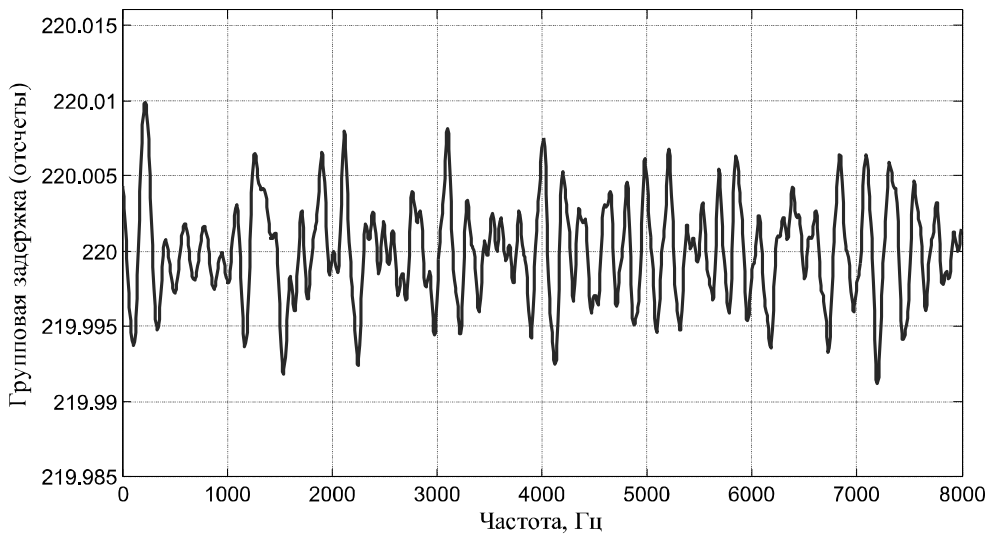


Рис. 5.19. Групповая задержка системы анализа/синтеза банка фильтров

5.2.3. Оценка эффективности компенсации сигнала АОС

Модель канала акустической обратной связи. Главный параметр, который влияет на акустическую обратную связь, – расстояние между динамиком и микрофоном слухового аппарата. Чем это расстояние меньше, тем больше будет эффект, вызываемый обратной связью. По этой причине слуховые аппараты категорий ITE (*in the-ear*) и ITC (*in-the-canal*), у которых данное расстояние минимально, больше всего подвержены влиянию

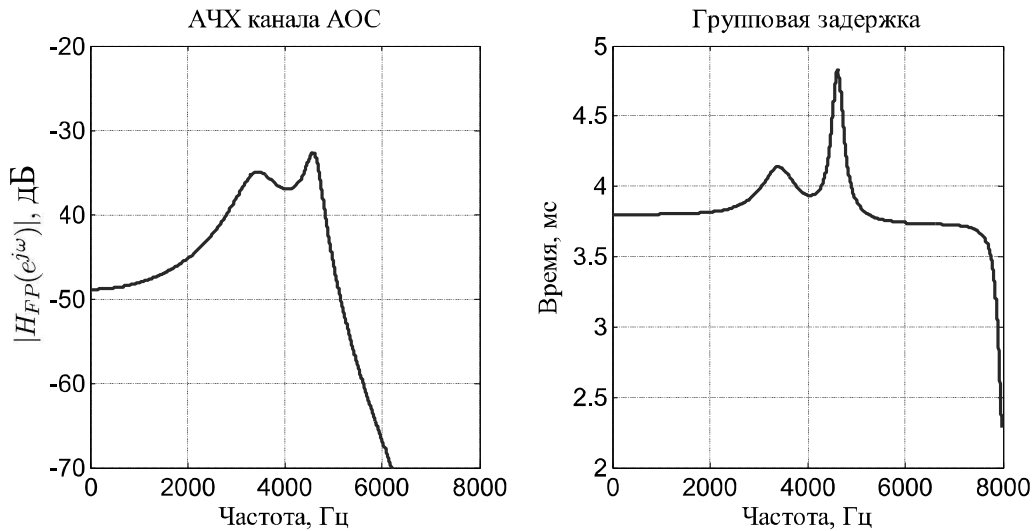


Рис. 5.20. Характеристики канала АОС: АЧХ и групповая задержка

обратной связи. Эмпирические исследования демонстрируют, что частотные характеристики каналов АОС различных слуховых аппаратов имеют много общих черт. В [44,128] показано, что канал АОС может быть смоделирован передаточной функцией

$$H_{FP}(s) = K e^{-sT} \frac{(s - z_1)(s - z_2)(s - z_2^*)}{(s - p_1)(s - p_1^*)(s - p_2)(s - p_2^*)}, \quad (5.5)$$

где K – коэффициент усиления;

T – задержка распространения сигнала в ушном канале между динамиком и микрофоном;

z_1, z_2, p_1, p_2 – параметры, определяющие канал АОС;

$s = \sigma + j\omega$ – комплексная переменная.

Для моделирования канала АОС слухового аппарата (типа ITE) использовались следующие параметры модели (5.5) [127]: $K = 1$; $T = 3,8$ мс; $z_1 = 1500000$; $z_2 = -40000 + 31400j$; $p_1 = -5000 + 22600j$; $p_2 = -500 + 33700j$. На рис. 5.20 изображены частотные характеристики моделируемого канала АОС.

Применяемые оценки. Для того чтобы оценить производительность системы компенсации эффекта АОС, воспользуемся параметрами SegSNR и SII, которые применялись для тестирования системы шумоподавления. Для перцептуальной оценки качества выходного сигнала использо-

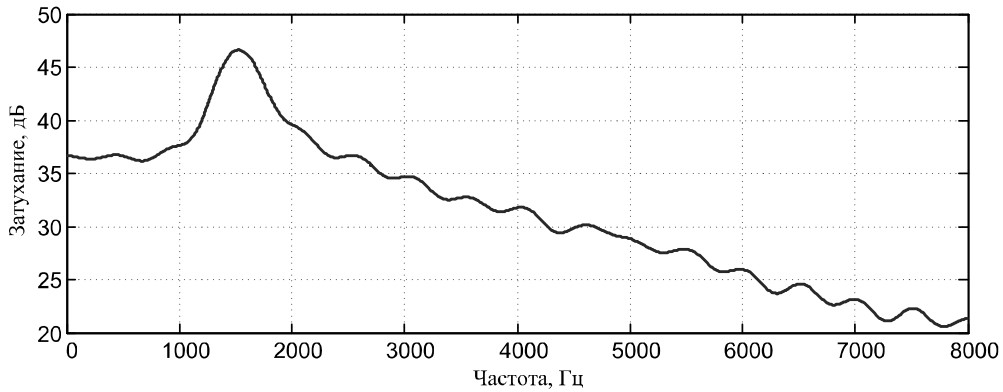


Рис. 5.21. Модель частотной характеристики слухового аппарата ($g_c = 46,5$ дБ)

вместо оценки PESQ (*perceptual evaluation of speech quality*) [129]. PESQ является объективным критерием качества канала связи, использующим разницу между спектрами громкости опорного и тестового сигналов для вычисления перцептуального качества тестового сигнала. Оценка PESQ колеблется в пределах от $-0,5$ до $4,5$. Наивысшая оценка показывает, что тестовый сигнал не содержит слышимых искажений и практически идентичен опорному. Оценки 4, 3 и 2 можно интерпретировать как качество «хорошее», «немного раздражающее», «раздражающее». Оценка между $-0,5$ и 1 свидетельствует о наличии неприемлемо большого количества шума и искажений в тестовом сигнале [6]. В качестве опорного сигнала в данном случае выступает выходной сигнал слухового аппарата в условиях отсутствия АОС, а в качестве тестового – сигнал слухового аппарата, в котором присутствует АОС и активна система компенсации.

Для моделирования прямого канала, т. е. обработки сигнала в слуховом аппарате, выбрана усредненная АЧХ (рис. 5.21), которая компенсирует типичное повреждение слуховой системы. Наибольшие потери происходят в районе 1,5 кГц, т. е. в том диапазоне частот, где речь имеет наибольшую информативность. Усиление выходного сигнала в СА регулируется при помощи коэффициента g_c .

Результаты экспериментов. Оценка производительности системы компенсации эффекта АОС производилась на речевом сигнале длительностью 5 минут (частота дискретизации 16 кГц). Сравнивались системы со следующими конфигурациями: без компенсации АОС, широкополосная

система компенсации эффекта АОС (см. рис. 5.13), субполосная система подавления эффекта АОС (см. рис. 5.15). Эксперименты проводились для различных порядков адаптивных фильтров ($L = 64, 128, 256$ и 512). Порядок субполосных адаптивных фильтров L_k в зависимости от L рассчитывался по следующему правилу:

$$L_k = \left\lceil \frac{L}{S_k} \right\rceil,$$

где S_k – коэффициент децимации/интерполяции в k -м канале банка фильтров, а $\lceil a \rceil$ обозначает наибольшее целое число, меньшее a .

Работа систем проверялась для различных значений коэффициента усиления g_c . В частности, рассматривались случаи $g_c = \{45,5; 46,5; 47,5; 48,5; 49,5\}$ дБ.³ В табл. 5.2–5.4 приведены основные результаты экспериментов (отметка НС в таблицах означает, что система нестабильна).

Результаты показывают, что применение субполосной техники адаптивной фильтрации позволяет более эффективно бороться с эффектом АОС. По ОСШ (см. табл. 5.2) субполосная система при больших коэффициентах усиления дает выигрыш примерно 1–2,5 дБ. Преимущества субполосной системы подавления выявляют как показатели перцептуального качества речи (PESQ), так и показатели индекса разборчивости речи (SII). Необходимо отметить и тот факт, что адаптивная система работает тем лучше, чем ближе порядок адаптивного фильтра к реальной длине канала АОС (в данном случае длина канала АОС равнялась приблизительно 70 интервалам дискретизации).

На рис. 5.22 показаны результаты работы системы на основе широкополосной компенсации эффекта АОС и системы на основе субполосной компенсации (для значений параметров $g_c = 47,5$ дБ, $L = 64$). Из приведенных спектрограмм видно, что скорость сходимости субполосной системы значительно выше, за счет чего уменьшается общий эффект от воздействия АОС. Предложенная система характеризуется повышенным порогом устойчивости и позволяет использовать более высокие коэффициенты усиления для обработки сигнала в прямом канале.

³При $g_c = 49,5$ дБ все системы начинают работать нестабильно, поэтому результаты экспериментов для этого случая в таблицах не отражены.

Таблица 5.2. Значения показателя SegSNR при различных коэффициентах усиления g_c и длине адаптивного фильтра L

Коэффициент усиления g_c , дБ	Без компенсации АОС	Широкополосная система компенсации АОС				Субполосная система компенсации АОС			
		64	128	256	512	64	128	256	512
45,5	4,06	4,50	4,40	4,45	4,394	5,06	5,03	4,97	4,71
46,5	0,34	2,09	2,05	2,09	2,032	3,31	3,31	3,24	2,89
47,5	НС	-1,47	-1,61	-1,39	-1,67	0,93	0,86	0,64	-0,23
48,5	~	НС	НС	НС	НС	-2,49	-2,87	-4,04	НС

Таблица 5.3. Значения PESQ при различных коэффициентах усиления g_c и длине адаптивного фильтра L

Коэффициент усиления g_c , дБ	Без компенсации АОС	Широкополосная система компенсации АОС				Субполосная система компенсации АОС			
		64	128	256	512	64	128	256	512
45,5	3,23	3,27	3,27	3,27	3,27	3,30	3,30	3,29	3,28
46,5	2,90	3,01	3,01	3,01	3,00	3,11	3,10	3,09	3,05
47,5	НС	2,62	2,62	2,67	2,67	2,87	2,86	2,84	2,78
48,5	—	НС	НС	НС	НС	2,13	2,16	2,26	НС

Таблица 5.4. Значения показателя SII при различных коэффициентах усиления g_c и длине адаптивного фильтра L

Коэффициент усиления g_c , дБ	Без компенсации АОС	Широкополосная система компенсации АОС				Субполосная система компенсации АОС			
		64	128	256	512	64	128	256	512
45,5	0,55	0,55	0,55	0,55	0,55	0,56	0,56	0,55	0,55
46,5	0,49	0,51	0,52	0,52	0,52	0,53	0,53	0,53	0,53
47,5	НС	0,46	0,46	0,46	0,45	0,49	0,49	0,48	0,47
48,5	—	НС	НС	НС	НС	0,40	0,39	0,37	НС

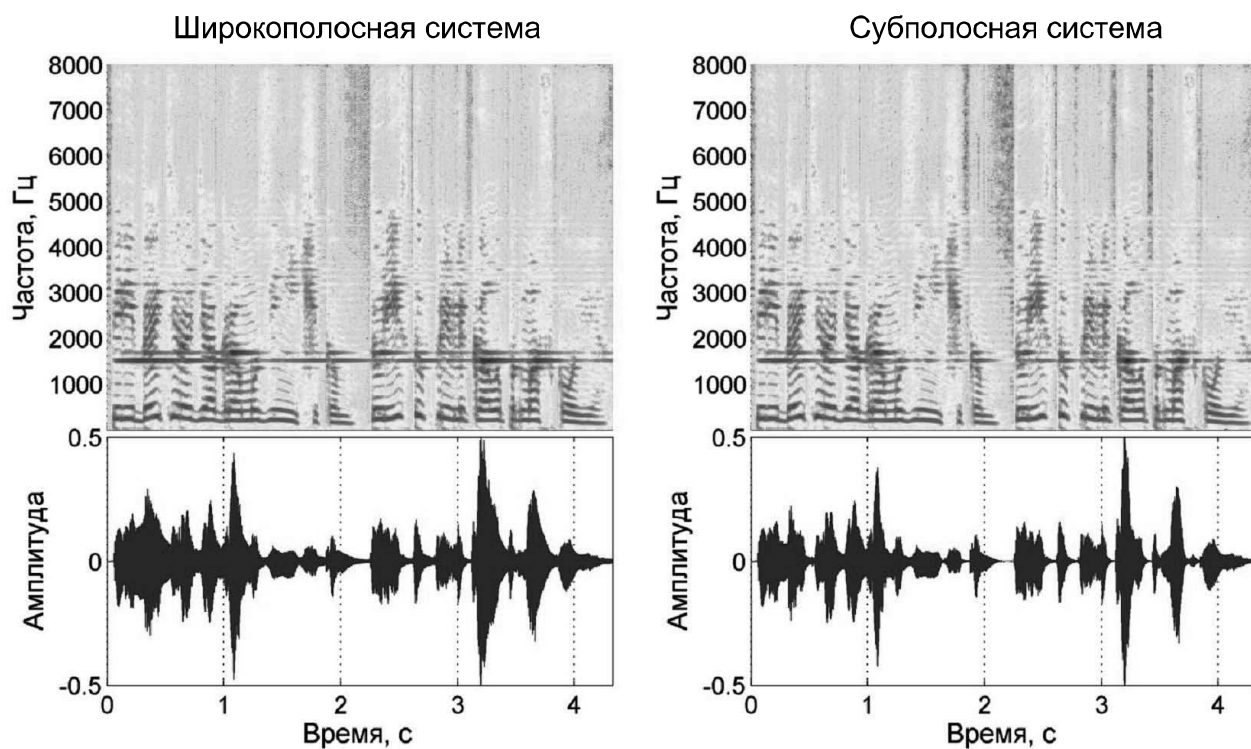


Рис. 5.22. Спектрограммы сигналов после обработки в широкополосной и субполосной системах

5.3. Основные результаты и выводы

Рассмотрены основные алгоритмы обработки сигнала в слуховом аппарате на основе неравнополосного банка фильтров.

1. Разработаны алгоритмы, выполняющие модификацию спектра сигнала в слуховом аппарате. В частности, разработан алгоритм, выполняющий функцию шумоподавления и основанный на перцептуально-мотивированном правиле спектрального взвешивания. Также проработан алгоритм, выполняющий компрессию динамического диапазона речевого сигнала для отображения широкого диапазона аудиосигнала в суженный диапазон остаточного слуха.

2. Предложен метод субполосной компенсации акустической обратной связи для слуховых аппаратов, основанный на использовании передискретизированного неравнополосного КМБФ, который не вносит дополнительной задержки в прямой путь распространения сигнала в слуховом аппарате. За счет разбиения сигнала на субполосы происходят более быстрая адаптация и более точная оценка компенсирующего сигнала обратной связи.

3. Выполнено тестирование предлагаемых алгоритмов на речевых сигналах. Полученные значения таких объективных показателей, как отношение сигнал-шум, перцептуальное качество канала связи, а также индекс разборчивости, показывают, что предложенные подходы к обработке сигнала в слуховом аппарате позволяют значительно улучшить качество выходного речевого сигнала.

ГЛАВА 6

РЕАЛИЗАЦИЯ СЛУХОВОГО АППАРАТА НА МОБИЛЬНОЙ ВЫЧИСЛИТЕЛЬНОЙ ПЛАТФОРМЕ

6.1. Преимущества и недостатки мобильных платформ для реализации слуховых аппаратов

В последнее время широкое распространение (в том числе и среди людей с патологиями слуха) получили распространение смартфоны – универсальные мультимедийные мобильные вычислительные платформы. Смартфоны имеют необходимые аппаратные средства для выполнения функций слухового аппарата: микрофон, АЦП, цифровой процессор, ЦАП, усилитель и динамики, причем возможно подключение внешнего микрофона и динамиков в виде специальной гарнитуры. Работа программных приложений с аппаратными ресурсами смартфона обеспечивается при помощи операционной системы, для которой производителем предоставляется пакет инструментов разработчика. Устойчивое развитие данных платформ, их поддержка и малая стоимость обусловлена популярностью смартфонов и многочисленностью их пользователей. Идея, положенная в основу данной работы, заключается в том, чтобы исследовать возможность создания карманного слухового аппарата на основе смартфона. В наиболее благоприятном случае, если аппаратные возможности смартфона соответствуют необходимым требованиям, создание слухового аппарата ограничится разработкой специального программного обеспечения, которое может использовать каждый пользователь с нарушениями слуха.

По сравнению с миниатюрными решениями карманный слуховой аппарат имеет следующие недостатки:

- карманный аппарат более заметен;
- в миниатюрном слуховом аппарате микрофон располагается в ухе и потому использует функциональные преимущества ушной раковины, в то время как в карманном слуховом аппарате такой возможности нет;

- аппарат не так удобен для ношения;

Однако, в силу своих конструктивных особенностей, карманные слуховые аппараты имеют важные преимущества, основные из которых перечислены ниже:

- большое расстояние между микрофоном и динамиком препятствует возникновению акустической обратной связи (АОС), что позволяет 1) использовать большое акустическое усиление; 2) использовать более простой алгоритм обработки (в миниатюрных слуховых аппаратах обязательна реализация схемы подавления АОС);
- большие физические размеры позволяют реализовать удобные функции управления для людей с плохой моторикой;
- возможно использование различных наушников в зависимости от патологии и предпочтений пациента (в том числе больших в виде накладных чашек, не приводящих к возникновению эффекта окклюзии);
- благодаря большим динамикам и большому ресурсу батарей можно создать наибольший уровень звукового давления и высокое качество звука;
- в отличие от внутриканальных СА, карманные не выходят из строя из-за попадания в них серы или влаги;
- при использовании динамика с костной проводимостью отсутствует механическая обратная связь.

Сверх того, благодаря тому, что слуховой аппарат реализован на популярной мобильной платформе, а не в виде отдельного узкоспециализированного устройства, для него характерны и некоторые дополнительные преимущества:

- большой доступный ресурс энергии и возможность использовать сложные алгоритмы обработки и высокую частоту дискретизации;
- большое распространение этих устройств, устойчивая поддержка и развитие которых осуществляется производителями благодаря многофункциональности и высокому спросу;
- программная гибкость – возможность программно реализовать произвольный алгоритм обработки сигналов;

- доступность средств разработчика и развитая система распространения программных продуктов для данных устройств;
- возможность применения алгоритмов коррекции слуха не только к акустическим сигналам, но и к аудио-видеозаписям, телефонным звонкам и т.д.;
- индивидуальная настройка аппарата для конкретного пациента в простых случаях не требует специального оборудования и квалификации и может выполняться самим пользователем;
- возможность удаленной настройки аппарата специалистом на основе диагноза пациента;
- наличие встроенной видеокамеры дает теоретическую возможность обработки комбинированной (речевой и видео) информации для повышения разборчивости речи;
- возможность использования различных гарнитур с разными характеристиками микрофонов и наушников, возможна разработка специальной гарнитуры с расположением микрофона в ушной раковине для использования звуковых характеристик наружного уха;
- заметность слухового аппарата не вызывает психологических неудобств, поскольку смартфон не ассоциируется у окружающих с какими бы то ни было слуховыми патологиями;
- для людей, уже постоянно пользующихся смартфонами, нет необходимости приобретать и носить отдельное устройство.

Целью настоящего исследования является реализация базовых функций слухового аппарата на мобильной платформе и экспериментальная оценка полученных результатов. Основной сложностью реализации является минимизация задержки обработки сигнала, поскольку задержка является определяющим фактором в системах коррекции слуха. Необходимо оценить минимально возможную задержку, вносимую смартфоном, и разработать схему коррекции с малой алгоритмической задержкой. В работе обсуждается классическая схема обработки сигнала, используемая в современных слуховых аппаратах, и предлагается альтернативная, учитывающая специфику вычислительной платформы.

6.2. Схема обработки сигнала для коррекции слуха

Основной функцией слухового аппарата является частотно-зависимое усиление входных сигналов для компенсации нарушений слуха [10]. Процесс обработки сигнала может включать вспомогательные функции, такие как подавление АОС, подавление шума и компрессия динамического диапазона. Все перечисленные функции, за исключением подавления АОС, могут быть выполнены путем субполосной декомпозиции и отдельным усилением каждого субполосного сигнала. На рис. 6.1 представлен пример соответствующей схемы обработки [6]. Входной звуковой сигнал $v(n)$ регистрируется микрофоном и раскладывается при помощи банка фильтров анализа (БФА) на субполосные сигналы с пониженной частотой дискретизации $u(m)$. Временные индексы исходного и субканальных сигналов обозначены n и m соответственно (контур, работающий на пониженной частоте, выделен жирными линиями на схеме).

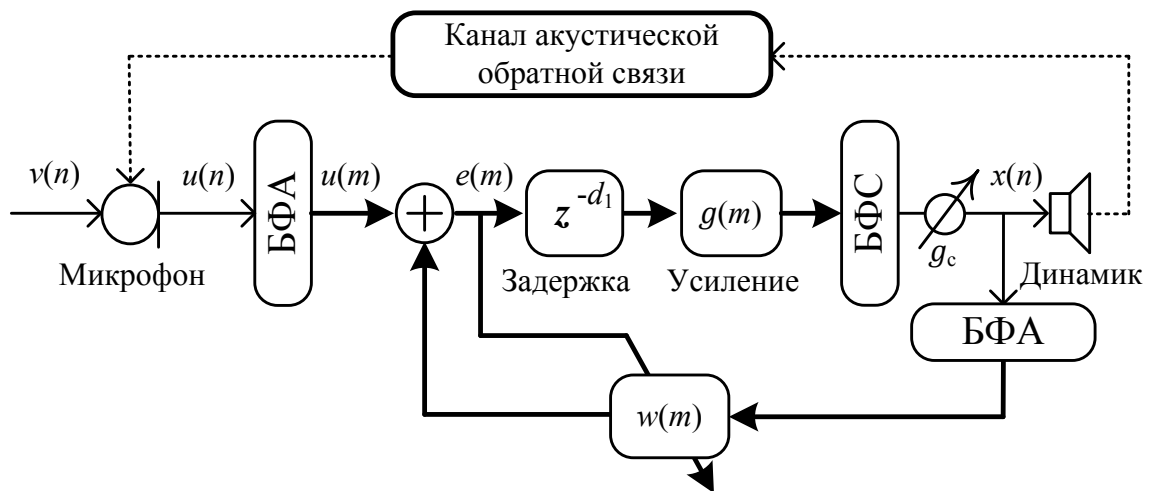


Рис. 6.1. Общая схема обработки сигнала в современном цифровом слуховом аппарате

Раздельное усиление каждого субполосного сигнала выполняется умножением на коэффициенты $g(m)$, которые вычисляются для каждого момента времени, исходя из уровня входного и выходного сигнала, энергии шума и частотного профиля потери слуха. Причем блоком усиления выполняется сразу несколько функций: 1) частотная коррекция потери чувствительности слуха; 2) компрессия (компенсация функции нелинейного усиления улитки); 3) шумоподавление на основе спектрального взвешивания.

вания. Обработанный широкополосный сигнал синтезируется при помощи банка фильтров синтеза (БФС). Широкополосный сигнал умножается на общий коэффициент усиления g_c , который обеспечивает комфортный уровень прослушивания (параметр g_c обычно регулируется пользователем при помощи внешнего регулятора).

В миниатюрных слуховых аппаратах максимальное доступное усиление сигнала ограничено появлением акустической обратной связи между микрофоном и динамиком. Для подавления АОС чаще всего используется схема адаптивной фильтрации. Коэффициенты адаптивного фильтра $w(m)$ постоянно обновляются таким образом, чтобы минимизировать энергию сигнала-остатка $e(m)$. Обновление коэффициентов выполняется во временной либо частотной области, используя какую-либо модификацию метода наименьших квадратов (LMS) [29, 127]. Задержка z^{-d_1} смещает сигналы друг относительно друга и влияет на оценку коэффициентов адаптивного фильтра. Большие значения d_1 обеспечивают более сильное подавление АОС и повышение возможной степени усиления сигнала [6]. Преимуществами схемы с субканальной декомпозицией сигнала являются относительно быстрая сходимость адаптивного фильтра и сокращение вычислительных ресурсов за счет понижения частоты дискретизации [130].

Компрессия сигнала выполняется с целью компенсации утерянной функции улитки внутреннего уха выполнять нелинейное усиление. Компрессия, применяемая для коррекции, должна иметь очень короткое время активации [10]. При использовании таких малых временных интервалов происходит быстрое изменение уровня сигнала, что приводит к появлению низкочастотных искажений. Дополнительным преимуществом приведенной выше схемы является возможность снизить уровень этих искажений, используя в низкочастотной части сигнала более продолжительное время активации.

Одной из наиболее важных характеристик слухового аппарата является задержка обработки сигнала. Временная задержка вызывает окрашивание звука собственного голоса пользователя слухового аппарата. Во время разговора голос быстро достигает улитки посредством костной проводимости и складывается с задержанным и усиленным звуком из слухового аппарата. Этот эффект становится более заметным при использовании

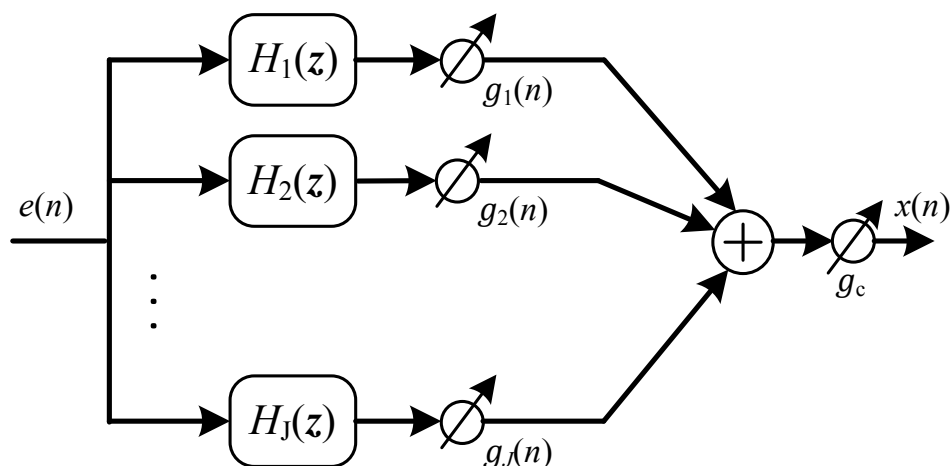


Рис. 6.2. Метод субканальной коррекции амплитудного спектра с малой задержкой

больших отверстий во вкладыше (вентов). В [131, 132] показано, что даже такие малые задержки обработки сигнала, как 4–8 мс ощущаются пациентом и снижают субъективное качество звука. В СА невозможно избежать задержек, обусловленных наличием АЦП и ЦАП, которые составляют от 0,4 до 2 мс в зависимости от реализации [133]. В условиях создания СА на базе мобильной платформы длительность неустраняемой аппаратной задержки будет существенно больше из-за внутренней аппаратной и программной реализации сигнальных конвейеров, поэтому в данном случае требуется схема обработки сигнала с минимально возможной алгоритмической задержкой. Общая алгоритмическая задержка классической схемы с субканальной декомпозицией, приведенной выше, является нежелательно большой [3, 134, 135] из-за последовательных элементов, включенных в путь прямого распространения сигнала: БФА, БФС и задержки z^{-d_1} , которая используется для стабилизации системы подавления АОС. В [6, 46] предлагаются схемы с субканальным усилением сигнала без БФС. Для этого используются полосовые фильтры (параметрические фильтры частотной компенсации [6] либо параметрические кохлеарные фильтры [46]), выходы которых суммируются (рис. 6.2).

В [46, 51] используется субканальная система подавления АОС, не вносящая дополнительных задержек в прямой канал. Перечисленные решения позволяют существенно сократить общую алгоритмическую задержку, однако отсутствие БФС делает невозможным понижение частоты дискретизации субполосных сигналов и требует больших вычислительных затрат.

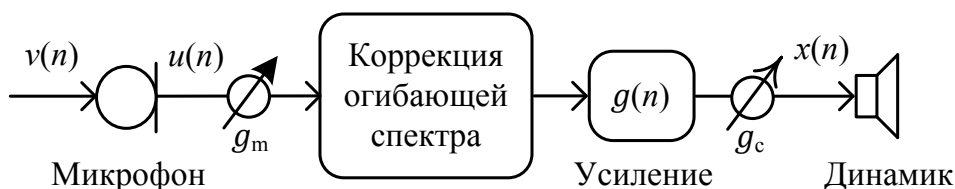


Рис. 6.3. Одноканальная схема обработки сигнала

Кроме того, полосовые фильтры с бесконечными импульсными характеристиками, используемые в [6, 46] имеют нелинейную фазо-частотную характеристику (ФЧХ), и для их согласования используется оптимизационная процедура, которая в зависимости от целевого амплитудного профиля подбирает субканальные коэффициенты усиления. Данная процедура также является достаточно затратной в вычислительном плане. Еще одной важной особенностью данной схемы является то, что обработка сигнала ведется отсчет за отсчетом, что предполагает использование специального аппаратного вычислителя. Существующие мобильные платформы позволяют обрабатывать сигнал в реальном времени только отдельными фреймами длиной не менее 6 мс. В этих условиях желательно, чтобы алгоритм обработки позволял использовать блочные операции. Учитывая перечисленные особенности, предлагается использовать широкополосную (без субполосной декомпозиции) схему коррекции слуха (рис. 6.3).

Схема использует широкополосный способ изменения коррекции огибающей спектра при помощи фильтра с конечной импульсной характеристикой (КИХ). Частотная характеристика фильтра формируется на основе аудиограммы пользователя таким образом, чтобы выполнять усиление частот, к которым наблюдается ослабление чувствительности. Для управления уровнем громкости используется два функциональных регулятора: регулятор чувствительности микрофона g_m и регулятор выходного уровня g_c . Оба регулятора являются внешними и доступны пользователю для оперативного использования в зависимости от акустической обстановки. Коэффициенты усиления $g(n)$ изменяются таким образом, чтобы обеспечивать компрессию сигнала в соответствии с заданным уровнем потери слуха.

Учитывая, что смартфон использует стереогарнитуру с двумя динамиками, можно использовать схему с бинауральной коррекцией слуха, выпол-

няющей обработку сигнала отдельно для левого и правого уха (рис. 6.4).

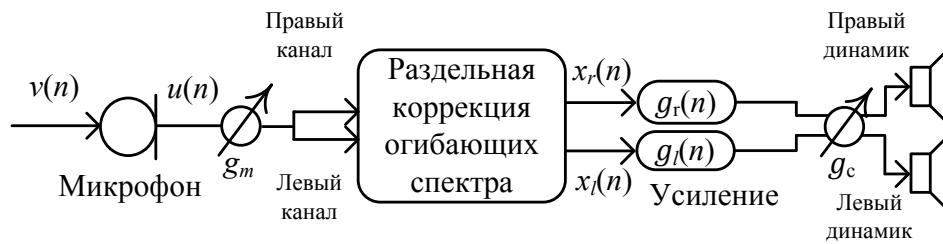


Рис. 6.4. Бинауральная схема обработки сигнала с одним микрофоном

Прямая реализация предлагаемой широкополосной схемы исключает использование шумоподавления на основе спектрального взвешивания, поскольку в ней не выполняется частотно-временное преобразование. Схема не содержит блока подавления АОС, поскольку при ее реализации на смартфоне предполагается использовать микрофон и динамик на большом расстоянии друг от друга. Благодаря тому, что предлагаемая широкополосная схема не содержит банка фильтров анализа, банка фильтров синтеза, а также стабилизирующего элемента задержки для подавления АОС, в ней достигается минимально возможная алгоритмическая задержка, равная групповой задержке фильтра коррекции спектральной огибающей.

6.3. Определение целевых параметров звукоусиления

Различают несколько составляющих тугоухости: ухудшение звуковосприятия, нарушения ощущения громкости и бинауральное снижение слуха [10]. Для оценки характеристик каждого из этих компонентов используют специальные методики и средства. Характер и степень ухудшения восприятия определяется при помощи аудиометрии, нарушения ощущений громкости оцениваются при помощи методов категоризации громкости, а бинауральные нарушения – путем тестирования различий разборчивости в каждом из ушей. В слуховом аппарате, реализованном на мобильной платформе с использованием предлагаемой схемы, можно выполнить компенсацию ухудшения звуковосприятия при помощи фильтра коррекции огибающей и компенсацию нарушения ощущения громкости (рекруитмента) при помощи элемента динамического усиления.

6.3.1. Аудиограмма

При подборе и настройке слухового аппарата аудиометрия дает самую первую и основную характеристику нарушения слуха. В результате аудиометрии формируется таблица минимальных уровней восприятия (*HTL – hearing threshold level*) тональных сигналов, называемая аудиограммой. Пороги восприятия определяются на частотах 125, 250, 500, 1000, 2000, 4000 и 8000 Гц и могут принимать значения от –10 до 120 дБ. В 1997 г. ВОЗ была утверждена единая Международная классификация степеней тугоухости, которая чаще всего используется на практике. Для определения степени тугоухости учитываются усредненные значения порогов звуковосприятия на четырех частотах: 500, 1000, 2000, 4000 Гц. Исходя из этой классификации, нормальным слухом считается слух при порогах слышимости на речевых частотах, не превышающих 10 дБ. Легкое нарушение слуха регистрируют при порогах слышимости 10–25 дБ. Потеря слуха свыше 90 дБ определяется как глухота.

Асимметричное нарушение слуха классифицируется по лучше слышащему уху. Выделяют IV степени тугоухости:

- I степень – снижение слуха в пределах 26–40 дБ (человек с такой потерей слуха с трудом распознает тихую речь и беседы, но справляется в тихой обстановке);
- II степень – 41–55 дБ (трудности в понимании беседы, особенно когда присутствует шум на заднем плане. Повышенная громкость необходима для ТВ и радио);
- III степень – 56–70 дБ (значительно задета чистота речи. Речь должна быть громкой, возможны трудности при групповой беседе);
- IV степень – 71–90 дБ (значительная потеря слуха – не слышит нормальную разговорную речь, трудности при распознавании даже громкой речи, способен понимать крик и преувеличенно четкую и громкую речь).

Аудиограмма дает возможность выполнить расчет необходимого (целевого) усиления и максимальной выходной мощности, обеспечивающих для пользователя максимально комфортный уровень громкости речевого сигнала при настройке обычного слухового аппарата. Снятие аудиограммы

пользователя выполняется в лабораторных условиях в помещении со звукоизоляцией и использованием специального оборудования, обеспечивающего генерацию и усиление тональных сигналов заданной частоты. Снятием аудиограммы руководит квалифицированный медицинский работник, создавая постепенно нарастающее звуковое давление в каждом ухе пациента. Как только звук становится различимым, пациент реагирует нажатием на сигнальную кнопку.

При помощи мультимедийной платформы можно автоматизировать весь процесс таким образом, чтобы пользователь мог самостоятельно выполнить аудиометрию. Для этого используются программные генераторы тональных сигналов и интерфейсная кнопка для реакции на превышение порога восприятия. Тональный сигнал синтезируется при помощи амплитудно-модулированного косинуса постоянной частоты:

$$s(n) = A(n) \cos\left(2\pi \frac{\omega n}{f_s}\right),$$

где $A(n)$ – изменяющаяся (нарастающая) амплитуда; ω – заданная частота в герцах и f_s – частота дискретизации сигнала. Скорость нарастания амплитуды определяется продолжительностью сигнала и доступным динамическим диапазоном ЦАП устройства. Нормированную амплитуду (единица соответствует максимально возможному значению) можно рассчитать по следующей формуле:

$$A(n) = 10^{(b-1)\left(\frac{n}{lf_s} - 1\right) \log_{10} 2},$$

где b – разрядность ЦАП и l – длительность сигнала в секундах. При таком способе расчета уровень сигнала в децибелах растет линейно от минимально возможного значения до максимального:

$$\text{level}(n) = 20 \log_{10} A(n) = 20(b-1) \left(\frac{n}{lf_s} - 1\right) \log_{10} 2.$$

Для снятия аудиограммы в лабораторных условиях используются специальные наушники с ровной амплитудно-частотной характеристикой (АЧХ). Для определения требуемого усиления в слуховом аппарате на каждой контрольной частоте учитываются частотные характеристики динамика слухового аппарата, используемого вента и ушного канала. В случае

выполнения измерений при помощи смартфона условия аудиометрии и эксплуатации являются одинаковыми, упрощенно можно считать, что неровности в АЧХ динамиков гарнитуры являются потерей слуха, и выполнить расчет целевых коэффициентов усиления без вносимых обычно поправок.

6.3.2. Расчет целевого усиления и коэффициентов корректирующего фильтра

Одним из наиболее известных способов расчета целевого усиления является правило полуусиления, согласно которому линейное усиление должно примерно равняться половине уровня потери слуха на каждой частоте для того, чтобы обеспечивать комфорт при прослушивании и речевую разборчивость [10]. Тем не менее, длительное время основным методом, использовавшемся в клинической практике, был сравнительный перебор, при котором пациент пробовал разные слуховые аппараты и выбирал тот, который обеспечивал наилучшие результаты. В исследовательской работе, опубликованной в 1975 году, был сделан вывод, что АЧХ слухового аппарата, определяющая уровень усиления в зависимости от частоты, существенно влияет на повышение разборчивости речи для людей с сенсоневральной тугоухостью. Это исследование инициировало постепенный процесс перехода от метода подбора к использованию директивных расчетных формул. Прежде чем появились нелинейные слуховые аппараты, методы расчетов частотно-зависимого линейного усиления представляли собой вариации вышеприведенного правила полуусиления.

Сегодня существует несколько классических способов расчета целевого усиления на основе аудиограммы: Berger [136], POGO (*Prescription Of Gain and Output* – предписание усиления и выхода) [137] и NAL-R (*National Acoustic Laboratories, Australia* – Национальные акустические лаборатории Австралии) [138]. Соответствующие расчетные формулы приведены в таблице 6.1 [10]. HTL_x обозначает измеренные значения порогов слышимости на частоте x .

Все три перечисленные метода расчета могут быть использованы для расчета целевого усиления в программной реализации слухового аппарата. Выбор наиболее подходящего из них делается либо самим пользователем

исходя из его предпочтений, либо специалистом на основе характера патологии и лабораторных измерений.

Таблица 6.1. Расчетные формулы целевого усиления

Частота, Гц	Berger	POGO	NAL-R
250		HTL250/2-10	X+0,31 HTL250-17
500	HTL500/2	HTL500/2-5	X+0,31 HTL500-8
750	HTL750/1,8	HTL750/2-3	X+0,31 HTL750-3
1000	HTL1k/1,6	HTL1k/2	X+0,31 HTL1k-1
1500	HTL1,5k/1,6	HTL1,5k/2	X+0,31 HTL1,5k-1
2000	HTL2k/1,5	HTL2k/2	X+0,31 HTL2k-1
3000	HTL3k/1,7	HTL3k/2	X+0,31 HTL3k-2
4000	HTL4k/1,9	HTL4k/2	X+0,31 HTL4k-2
6000	HTL6k/2	HTL6k/2	X+0,31 HTL6k-2

Для метода NAL-R коэффициент X рассчитывается следующим образом:

$$X = 0,05(\text{HTL500} + \text{HTL1k} + \text{HTL2k}).$$

Для тяжелой и глубокой степени потери слуха метод NAL-R был переработан. Модифицированный вариант метода, называемый NAL-RP [139], дает большее усиление и более равномерную частотную характеристику, чем NAL-R. Коэффициент X вычисляется по формуле

$$X = \begin{cases} 0,05(\text{HTL500} + \text{HTL1k} + \text{HTL2k}), \\ \text{если } (\text{HTL500} + \text{HTL1k} + \text{HTL2k}) < 180 \text{ дБ;} \\ 0,116(\text{HTL500} + \text{HTL1k} + \text{HTL2k}), \\ \text{если } (\text{HTL500} + \text{HTL1k} + \text{HTL2k}) > 180 \text{ дБ.} \end{cases}$$

Если порог слышимости на частоте 2 кГц превышает 95дБ, к полученным значениям усиления нужно прибавить поправочные коэффициенты, приведенные в таблице 6.2.

В соответствии с полученными значениями целевого усиления рассчитываются коэффициенты фильтра коррекции спектральной огибающей.

Таблица 6.2. Поправочные коэффициенты при тяжелых формах тугоухости

Пороги слышимости на частоте 2 кГц	250 Гц	500 Гц	750 Гц	1000 Гц	1500 Гц	2000 Гц	3000 Гц	4000 Гц	6000 Гц
95	4	3	1	0	-1	-2	-2	-2	-2
100	6	4	2	0	-2	-3	-3	-3	-3
105	8	5	2	0	-2	-3	-5	-5	-5
110	11	7	3	0	-3	-6	-6	-6	-6
115	13	8	4	0	-4	-8	-8	-8	-8
120	15	-	4	0	-5	-9	-9	-9	-9

В данной реализации слухового аппарата используются фильтры с линейной ФЧХ, что позволяет достигать равной временной задержки на всех частотах обрабатываемого сигнала, вне зависимости от целевых значений усиления. Данное свойство особенно важно в бинауральном режиме коррекции слуха, поскольку различия в групповых задержках фильтров для правого и левого каналов вызывают фазовые смещения и, как следствие, искажения пространственного восприятия звука. Фильтры с линейной ФЧХ удобно синтезировать при помощи оконного метода [140]. Алгоритм получения коэффициентов фильтра кратко приведен ниже.

- 1) Целевая частотная характеристика фильтра формируется путем интерполяции целевых значений усиления в точках, соответствующих частотным отсчетам. Фаза всех частотных компонент принимается нулевой.
- 2) Вычисляется обратное преобразование Фурье целевой частотной характеристики.
- 3) Выполняется циклический сдвиг полученного вектора значений на половину числа его элементов.
- 4) Окончательная импульсная характеристика фильтра формируется путем умножения полученных значений на оконную функцию.

Число коэффициентов фильтра выбирается исходя из частоты дискретизации сигнала и допустимой групповой задержки.

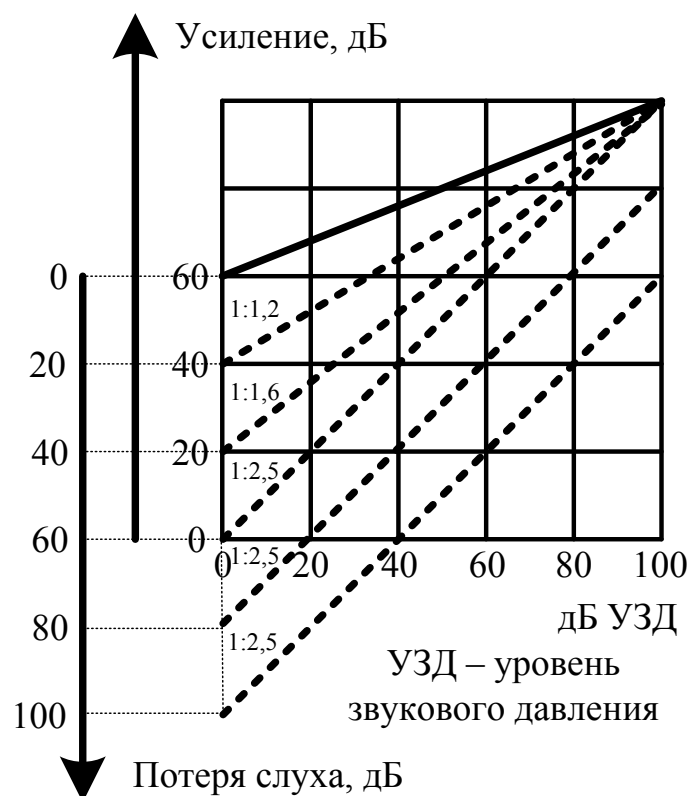


Рис. 6.5. Усиление в улитке

6.3.3. Компенсация нарушения ощущения громкости

Установлено, что тихие звуки усиливаются здоровой улиткой на 50–60 дБ, а громкие не усиливаются. Изменение коэффициента усиления физиологически обеспечивается наружными волосковыми клетками. При потере слуха более чем на 60 дБ волосковые клетки теряют свои регулятивные возможности и характеристика улиткового усилителя превращается в линейную – рис. 6.5 [10]. В этом случае требуется компрессия с коэффициентом 2,5:1.

Слышимый динамический диапазон пациента сужается, и пациент может относительно хорошо воспринимать громкие звуки по сравнению с тихими. Другими словами, при увеличении уровня сигнала восприятие громкости при тугоухости приближается к нормальному. Это требует некоторого увеличения громкости, которое называется *рекруитментом*. Основным назначением компрессора является усиление тихих звуков, однако в то же время предохранение чрезмерного усиления громких звуков. Общая схема, приведенная на рис. 6.6, показывает функциональную организацию простого компрессора. Сперва измеряется текущий уровень сигнала, а за-

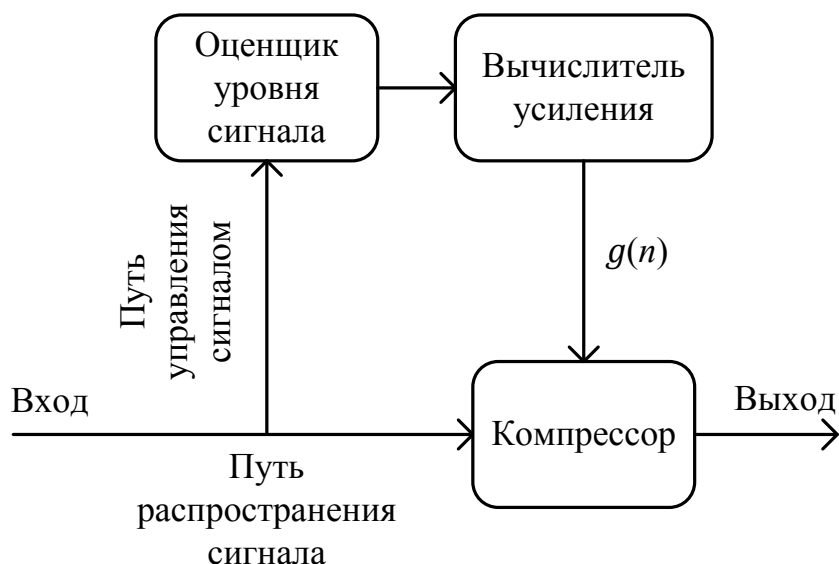


Рис. 6.6. Простая схема компрессии сигнала

тем вычисляются требуемые коэффициенты усиления для данного уровня $g(n)$, исходя из степени потери слуха.

Компрессор, схема которого приведена на рис. 6.6, называется компрессором с прямой связью, поскольку в нем выполняется оценка уровня входного сигнала перед выполнением компрессии. Если же выполняется оценка уровня выходного сигнала после компрессии, то такая схема называется схемой с обратной связью. В слуховом аппарате преимущественно используются компрессоры с прямой связью.

6.3.4. Временные параметры компрессора динамического диапазона

Динамические характеристики компрессора определяются скоростью реакции оценщика уровня сигнала на изменения уровня сигнала. Для того чтобы избежать появления артефактов, изменения коэффициентов усиления не должны быть резкими. Это достигается сглаживанием действительной амплитудной огибающей сигнала. Системы компрессии с сильным сглаживанием могут использоваться при частой смене акустической обстановки либо для выравнивания общего уровня прослушивания. В этом случае система реагирует медленно и иногда называется системой автоматического регулирования усиления. Системы с коротким времени реакции имеют своей целью регулировку усиления перед каждой новой фонемой

или слогом обрабатываемой речи, в этом случае компрессор часто называют фонемным или слоговым. Ни один оценщик уровня не является идеальным, всегда существует некоторая задержка между реальным уровнем сигнала и его оценкой. Быстрота реакции оценщика уровня определяется временными константами.

Оценка уровня сигнала $x(n)$ в заданный момент времени t_n обозначается $\Phi(t_n)$ и обновляется с заданным фиксированным временным интервалом $\Delta t = t_n - t_{n-1}$ [141]. Обычно $\Phi(t_n)$ вычисляется как взвешенное среднее между его предыдущим значением $\Phi(t_{n-1})$ и средним квадратом отсчетов сигнала $\bar{x}_{[t_{n-1}, t_n]}^2$, находящихся в интервале между t_{n-1} и t_n :

$$\Phi(t_n) = a\Phi(t_{n-1}) + (1 - a)\bar{x}_{[t_{n-1}, t_n]}^2,$$

где параметр a находится в интервале от 0 до 1. Выбор константы a определяет скорость реакции оценщика уровня. Если a близко к 1, оценка уровня меняется очень медленно, в то время как при приближении a к 0 время реакции уменьшается.

Еще одна временная константа, обозначенная как τ , используется для настройки параметров компрессора и принимает различные значения в зависимости от того, повышается или понижается уровень сигнала. Если уровень сигнала повышается, то она называется временем срабатывания (*attack time*), если понижается, то временем восстановления (*release time*). В соответствии со стандартом IEC 60118-2 время срабатывания, обозначаемое как α , задается как время, необходимое для того, чтобы выход достиг разницы в 2 дБ от окончательного значения при резком увеличении уровня звукового давления (УЗД) сигнала от 55 до 80 дБ. Время восстановления, обозначаемое как β , является временем, необходимым для того, чтобы достичь разницы в 2 дБ от окончательного значения при резком изменении уровня сигнала от 80 дБ УЗД до 55 дБ УЗД. В обоих случаях предполагается, что выходной уровень достиг устойчивого значения прежде резкого изменения уровня.

Допустим, что сигнал изменяется от x_1 до x_2 в момент времени t_1 , т.е. $x(t_n) = x_1$ для $n \leq 0$ и $x(t_n) = x_2$ для $n > 0$, и $\Phi_{t_0} = x_1^2$. Тогда

$$\Phi(t_n) = a^n(x_1^2 - x_2^2) + x_2^2.$$

Допустим, что требуемый выходной уровень достигается в момент времени t_{n_τ} , где τ означает временную константу (время срабатывания либо восстановления). Тогда τ можно вычислить из заданного значения веса a :

$$\tau = n_\tau \Delta t = \frac{\ln(x_2^2 - \Phi_{t_{n_\tau}}) - \ln(x_2^2 - x_1^2)}{\ln a} \Delta t.$$

Искомый весовой коэффициент вычисляется из временной константы τ следующим образом:

$$a = \left(\frac{x_2^2 - \Phi_{t_{n_\tau}}}{x_2^2 - x_1^2} \right)^{\Delta t / \tau}.$$

При вычислении весового коэффициента для заданного времени срабатывания значения x_1^2 и x_2^2 устанавливаются равными значениям соответствующим 55 дБ УЗД и 80 дБ УЗД соответственно; при вычислении весового коэффициента для заданного времени восстановления значения x_1^2 и x_2^2 устанавливаются равными значениям соответствующим 80 дБ УЗД и 55 дБ УЗД соответственно. Преимуществом подхода использования временной константы является то, что он остается универсальным вне зависимости от частоты дискретизации сигнала и от частоты обновлений уровня сигнала.

Выбор временной константы в компрессоре слухового аппарата является дискуссионным вопросом. Основной целью быстрой компрессии является регулировка усиления перед каждым звуком речи для того, чтобы тихие фонемы (согласные) получали более высокий уровень усиления, чем громкие. В опубликованных научных результатах не было показано, что компрессия (медленная либо быстрая) повышает разборчивость по сравнению с линейной системой усиления, обеспечивающей постоянно комфортный для прослушивания уровень сигнала. Однако в том случае, когда речь имеет широкий динамический диапазон, компрессия имеет заметное преимущество. Некоторые исследования показали, что пациенты со слуховыми патологиями предпочитают качество звука, обеспечиваемое системой компрессии с медленной реакцией, особенно в шумовой обстановке [141]. Одна из проблем медленной компрессии это недостаточное усиление тихих звуков, следующих сразу за громкими, поскольку компрессору нужно продолжительное время для регулировки усиления. Это может вызывать у пользователя слухового аппарата пропуски фрагментов речи после гром-

ких звуков. Напротив, проблемой при использовании быстрой компрессии является то, что тихий шум становится слышимым на тихих фрагментах либо паузах речи. Показано, что при наличии статического фонового шума быстрая компрессия предпочтительнее медленной. Согласно данным, приведенным в [10], время активации должно быть коротким для того, чтобы обеспечивать слоговую компрессию (время срабатывания <10 мс и время восстановления <50 мс).

6.4. Реализация слухового аппарата на мобильной платформе

6.4.1. Схема включения приложения

В качестве мобильной платформы используется смартфон компании Apple семейства iPhone5. Для реализации слухового аппарата необходимо выполнять обработку в реальном времени с малой алгоритмической задержкой. Программное приложение должно установить потоковую передачу цифрового сигнала от микрофона к динамику. Внутренняя организация функциональных блоков iPhone для обработки аудио позволяет применить алгоритм коррекции слуха к сигналу, формируемому на выходе микшера, как показано на рис. 6.7. Микшер выполняет синхронизацию и захват сигналов из микрофонной линии и программных приложений системы. В результате можно выполнять коррекцию не только микрофонного сигнала, но и сигналов различных программных приложений смартфона, таких как проигрыватель аудио и видео.

Приложение для коррекции слуха имеет два режима: режим аудиометрии и режим коррекции. В режиме аудиометрии приложение измеряет пороги слышимости путем генерации тональных сигналов. В режиме коррекции выполняется обработка сигнала с учетом полученных порогов. Пользователь может подобрать параметры для достижения наилучшей разборчивости.

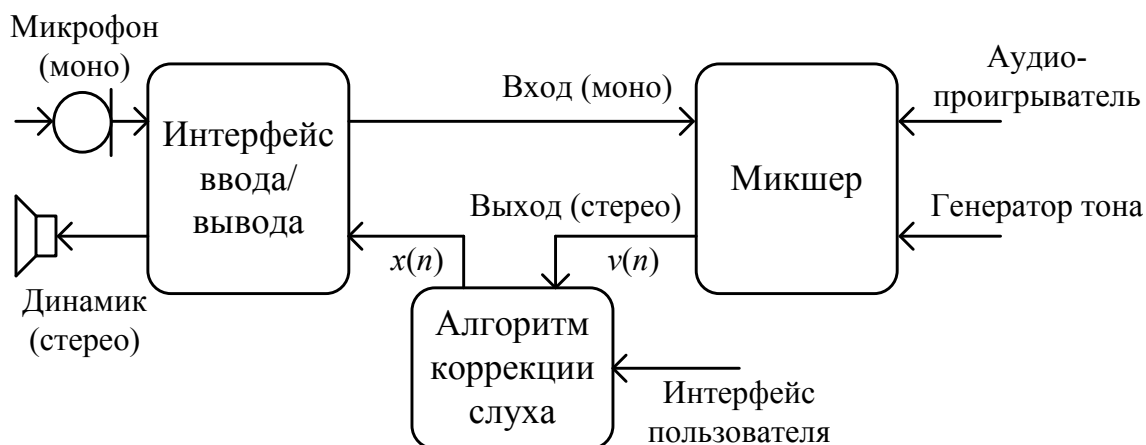


Рис. 6.7. Схема включения приложения коррекции слуха в звуковую систему iPhone

6.4.2. Полоса пропускания

Полоса пропускания в современных слуховых аппаратах не превышает 10 кГц, поскольку используемая частота дискретизации составляет 16–20 кГц и существенная часть слышимого звукового спектра не обрабатывается. Ограничение частоты вызвано тем, что используемые в слуховых аппаратах специализированные цифровые процессоры обрабатывают сигнал с частотой, равной или пропорциональной частоте дискретизации, и ее повышение приводит к увеличению потребляемой мощности и сокращению срока службы батарейки. Потому перед разработчиками слуховых аппаратов всегда стоит выбор: расширить полосу пропускания либо увеличить срок службы элементов питания. Учитывая, что платформа iPhone имеет перезаряжаемый элемент питания и ориентирована на мультимедийные приложения, возможно использование частоты дискретизации 44,1 кГц, обеспечивающей наилучшее качество звука.

6.4.3. Алгоритм обработки сигнала в слуховом аппарате на мобильной платформе

Вычислительная платформа iPhone позволяет обрабатывать входной сигнал отдельными кадрами по R отсчетов. Размер кадра как правило колеблется в пределах от 128 до 1024 отсчетов и выбирается равным степени числа 2. Для реализации функции слухового аппарата с целью уменьшения задержки лучше выбирать размер кадра равным 128 или 256. Обработка

сигнала заключается в выполнении двух основных процедур: 1) коррекция огибающей спектра (на основе аудиограммы пользователя) и 2) динамическая компрессия сигнала с целью компенсации нарушения ощущения громкости.

Коррекция огибающей спектра может быть выполнена при помощи КИХ-фильтра с фиксированной частотной характеристикой. Фильтр синтезируется заранее по индивидуальной аудиограмме пользователя с использованием правил расчета целевого усиления POGO, NAL-R или Berger [10]. Для уменьшения вычислительной сложности фильтрацию необходимо выполнять в частотной области. Для этой цели предлагается использовать метод перекрытия с суммированием [140]. Схема алгоритма обработки сигнала для реализации функции слухового аппарата показан на рис. 6.8. На вход поступает дискретизированный сигнал $x[n]$, который разбивается на кадры $x^{(m)}[r]$:

$$x^{(m)}[r] = x[mR + r],$$

где m – номер кадра;

$r = 1, \dots, R$ – индекс отсчета внутри кадра.

Для выполнения линейной фильтрации методом перекрытия с суммированием входной кадр расширяется последовательностью из R нулевых отсчетов. К полученной последовательности применяется алгоритм быстрого преобразования Фурье (БПФ) для перевода сигнала из временной области в частотную. Результат преобразования умножается на заранее рассчитанную частотную характеристику КИХ фильтра-корректора. Для перевода сигнала во временную область используется алгоритм обратного быстрого преобразования Фурье (ОБПФ). Результатом преобразования является частичная свертка кадра входного сигнала $x^{(m)}[r]$ с импульсной характеристикой фильтра-корректора. Частичная свертка имеет два участка $y_1^{(m)}[r]$ и $y_2^{(m)}[r]$ (см. рис. 6.8). Текущий результат фильтрации, который соответствует кадру m , формируется путем суммирования:

$$y^{(m)}[r] = y_1^{(m)}[r] + y_2^{(m-1)}[r].$$

Следующим этапом обработки сигнала является компрессия динамического диапазона (КДД) с целью компенсации нарушения ощущения

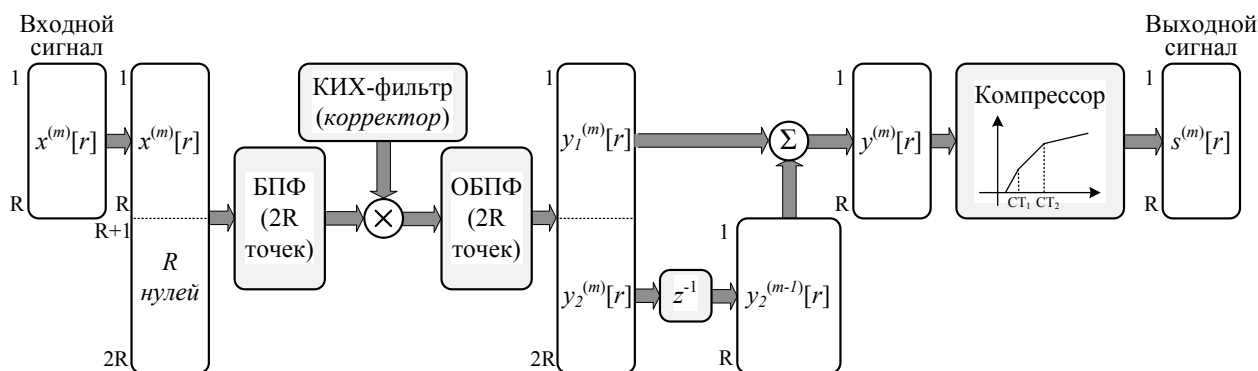


Рис. 6.8. Алгоритм обработки сигнала для реализации функции слухового аппарата

громкости. Главной задачей КДД является автоматическое регулирование коэффициента усиления сигнала ($g(n)$ на рис. 6.3) в зависимости от уровня мощности сигнала. Характер КДД определяется функцией вход/выход компрессора, которая показывает, как должен изменяться уровень выходного сигнала при изменении уровня входного сигнала. В работе используется компрессор широкого динамического диапазона (*WDRC – wide dynamic range compression*), характеристика вход/выход которого показана на рис. 6.9. Динамический диапазон входного сигнала условно разбит на три неперекрывающиеся области: «шум», «тихие звуки» и «речь». Под «шумом» понимается собственный шум внутренней схемы платформы iPhone, которая отвечает за прием и дискретизацию, и квантования входного акустического сигнала. Это сигнал малой интенсивности, который при прохождении компрессора не усиливается. Выше уровня «шума» находятся «тихие звуки». Сигналы, относящиеся к этой категории, плохо различимы людьми, страдающими тугоухостью, поэтому при их обнаружении компрессор КДД начинает работать в режиме расширения (*expansion*), который характеризуется коэффициентом расширения XR . Если Δx это изменение уровня входного сигнала, а Δy – выходного, то

$$XR = \Delta x / \Delta y.$$

Главной особенностью режима расширения КДД является то, что коэффициент XR всегда меньше единицы. Например, для коэффициента расширения $XR = 0,5$ изменение входного уровня сигнала на 10 дБ приведет к изменению выходного уровня на 20 дБ.

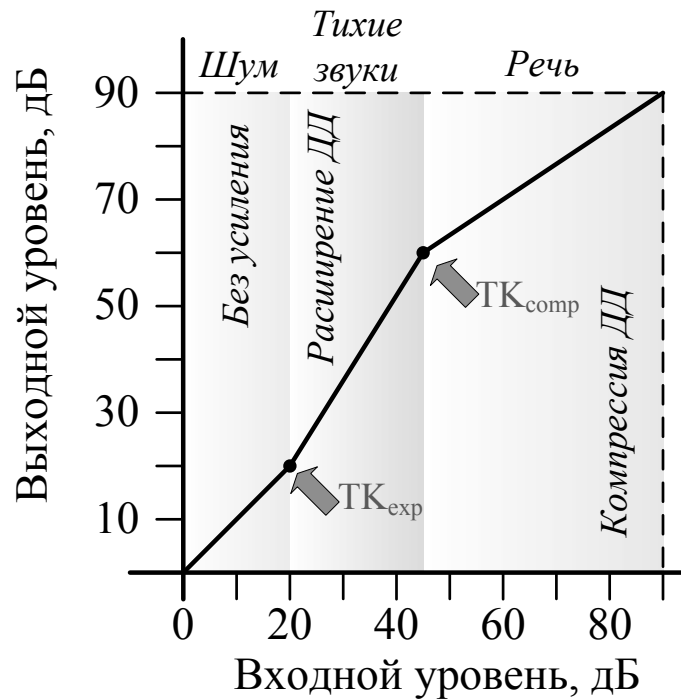


Рис. 6.9. Характеристика вход/выход КДД

Если уровень входного сигнала попадает в категорию «речь» (рис. 6.9), то активизируется режим компрессии динамического диапазона, который характеризуется коэффициентом компрессии

$$CR = \Delta x / \Delta y,$$

который всегда больше единицы. Например, при коэффициенте $CR = 2$, изменение входного уровня сигнала на 10 дБ приведет к изменению выходного уровня на 5 дБ.

Таким образом, характеристика вход/выход компрессора имеет вид кусочно-линейной кривой, имеющей характерные точки: TK_{exp} – переход к режиму расширения и TK_{comp} – переход в режим компрессии. Кроме характеристики вход/выход компрессор имеет параметры времени срабатывания (атаки) и восстановления (их смысл пояснен выше).

В соответствии с рис. 6.8 на вход компрессора поступает сигнал $y^{(m)}[r]$ от фильтра-корректора. Задача компрессора состоит в расчете линейной функции усиления $g[r]$ для формирования выходного сигнала:

$$s^{(m)}[r] = y^{(m)}[r] \cdot g[r], \quad r = 1, \dots, R.$$

Поскольку на функцию $g[r]$ наложено ограничение линейности, то достаточно определить значения $g[1]$ и $g[R]$, а остальные значения находятся

путем линейной интерполяции. Ниже приведен псевдокод для алгоритма определения $g[1]$ и $g[R]$:

```

1:  $P_{\text{start}} = 10 \log_{10}(p)$ 
2: for  $r = 1, 2 \dots, R$  do
3:    $p_{\text{inst}} = (y^{(m)}[r])^2$ 
4:   if ( $p_{\text{inst}} > p$ ) then
5:      $p = \alpha p + (1 - \alpha)p_{\text{inst}}$ 
6:   else
7:      $p = \beta p + (1 - \beta)p_{\text{inst}}$ 
8:   end if
9: end for
10:  $P_{\text{end}} = 10 \log_{10}(p)$ 
11:  $P_{\text{start}}^{\text{out}} = \text{IO\_Func}(P_{\text{start}})$ 
12:  $P_{\text{end}}^{\text{out}} = \text{IO\_Func}(P_{\text{end}})$ 
13:  $g[1] = 10^{(P_{\text{start}}^{\text{out}} - P_{\text{start}})/20}$ 
14:  $g[R] = 10^{(P_{\text{end}}^{\text{out}} - P_{\text{end}})/20}$ 

```

где α – параметр усреднения, зависящий от времени атаки; β – параметр усреднения, зависящий от времени восстановления.

В приведенной программе переменная p хранит текущее (среднее) значение уровня мощности сигнала. Для первого кадра сигнала ($m = 0$) p присваивается значение 2^{-10} , в последующем значение p сохраняется от кадра к кадру. Таким образом, P_{start} для m -го кадра равно значению P_{end} ($m-1$)-го кадра. Текущее значение мощности сигнала находится путем экспоненциального усреднения. Параметр экспоненциального усреднения выбирается в зависимости от того, что происходит: нарастание или спад сигнала. Через $\text{IO_Func}()$ в программе обозначена функция вход/выход КДД (рис. 6.9).

На рис. 6.10 показан пример обработки тестовых тональных сигналов в компрессоре динамического диапазона. Пример иллюстрирует характер изменения коэффициента усиления в компрессоре в зависимости от уровня входного сигнала, а также поясняет значения понятий времени атаки и времени восстановления.

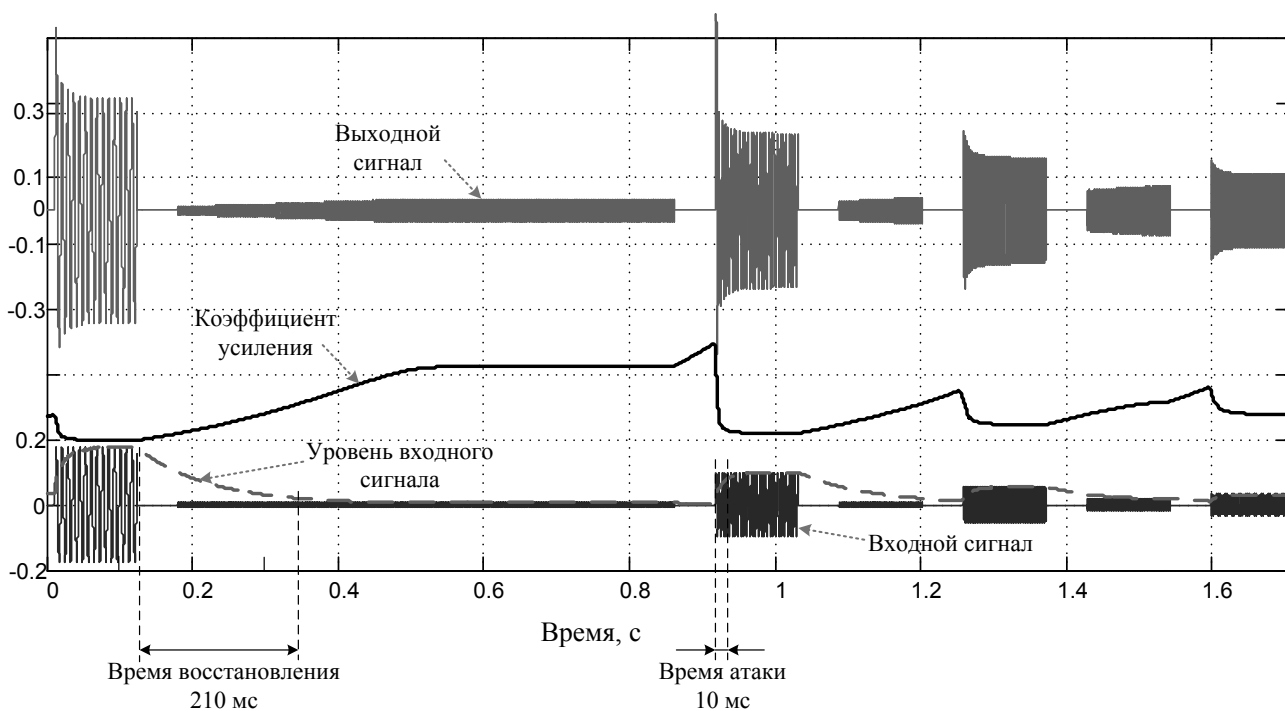


Рис. 6.10. Обработка тестовых тональных сигналов компрессором динамического диапазона

Пример обработки речевого сигнала разработанным алгоритмом показан на рис. 6.11. Можно видеть, что совместное использование фильтра-корректора и компрессора динамического диапазона позволяет выполнить уровне- и частотно-зависимую обработку сигнала. Следует отметить обработку фрагмента речевого сигнала, отвечающего шипящему звуку «с». Значительное усиление этого звука обусловлено тем, что его основная энергия лежит в верхней части частотного диапазона, который плохо воспринимается людьми, страдающими тугоухостью. По тем же причинам значительное усиление получил взрывной звук «п».

6.4.4. Задержка обработки сигнала

Задержка обработки сигнала складывается из групповой задержки корректирующего фильтра и задержки ввода/вывода аудиосистемы вычислительной платформы. Учитывая, что используется фильтр с линейной ФЧХ, групповая задержка корректирующего фильтра равна половине импульсной характеристики. Для формирования фильтра с удовлетворительной аппроксимацией целевой частотной характеристики нужно использовать 128–256 коэффициентов, что при частоте дискретизации 44,1 кГц со-

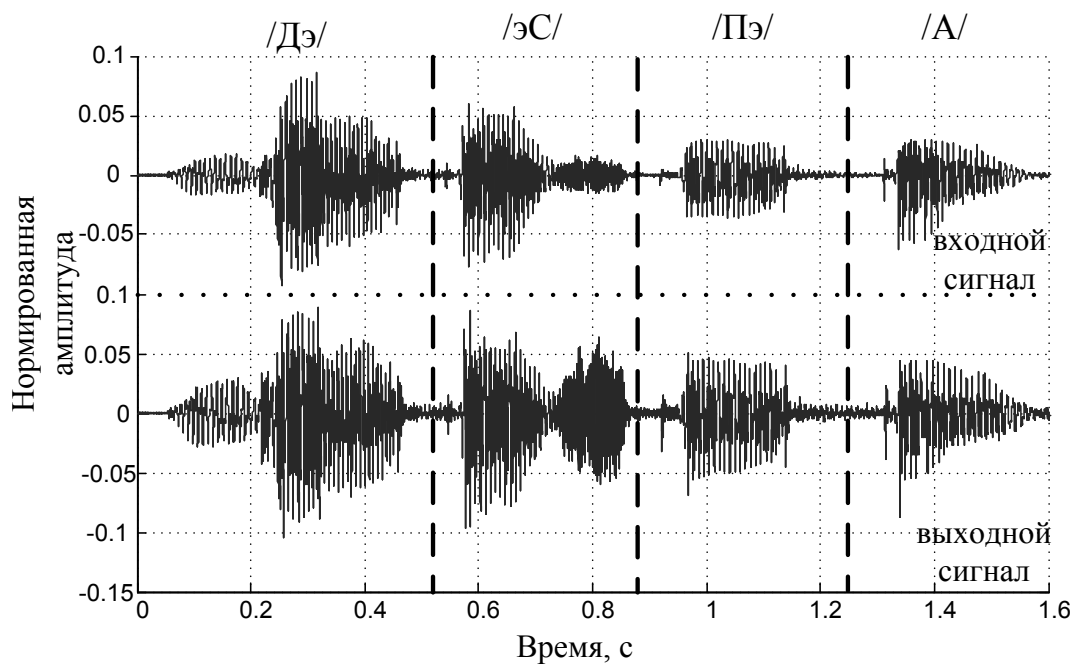


Рис. 6.11. Обработка речи в слуховом аппарате

ответствует групповой задержке 1,45–2,9 мс. Организация ввода/вывода включает неустраняемые задержки ЦАП и АЦП, которые составляют примерно 2–4 мс, и задержки буфера сигнала. В отличие от слуховых аппаратов, использующих специальные вычислители, на платформе iPhone нет возможности выполнять захват сигнала короткими фреймами в несколько отсчетов, поскольку это приводит к неоправданно высоким затратам на обслуживание приложения со стороны операционной системы. В зависимости от модели устройства приемлемым является ввод и вывод сигнала блоками по 128–256 отсчетов, что соответствует задержке 2,9–5,8 мс. Таким образом, общая задержка обработки сигнала при реализации слухового аппарата на платформе iPhone может составлять от 6,3 до 15,7 мс.

6.5. Результаты экспериментов

6.5.1. Конфигурация устройства

Для экспериментов использовалось устройство iPhone5 со штатной гарнитурой. Операционная система iOS 7.0 позволяет использовать встроенный микрофон, находящийся на панели устройства, одновременно с наушниками гарнитур. Экспериментальным путем установлено, что встро-

енный микрофон обеспечивает более высокое соотношение сигнал/шум и потому предпочтительнее для приложения. Более того, большое расстояние между наушниками гарнитуры и встроенным микрофоном исключает появление АОС. Исходя из перечисленных особенностей во всех экспериментах, проведенных в рамках этой работы, использовался встроенный микрофон устройства. Использовалась высокая частота дискретизации сигнала 44.1 кГц и разрядность 16 бит.

В процессе аудиометрии пользователю подавались сигналы заданной частоты в следующей последовательности: 1000, 2000, 4000, 6000, 125, 250, 500 Гц. Сигналы синтезировались с нарастающей амплитудой от минимально возможного значения до максимального на протяжении 15 секунд.

6.5.2. Полное акустическое усиление

Акустическим усилением называется разность между выходным уровнем звукового давления, создаваемого слуховым аппаратом в имитаторе уха, и входным уровнем звукового давления, измеренным в тестовой точке. Полным акустическим усилением называется акустическое усиление слухового аппарата при линейном входе/выходе, максимальном положении регуляторов слухового аппарата.

Лабораторные измерения показали, что аудиосистема iPhone5 со стандартной гарнитурой позволяет создавать полное акустическое усиление до 30 дБ – рис. 6.12.

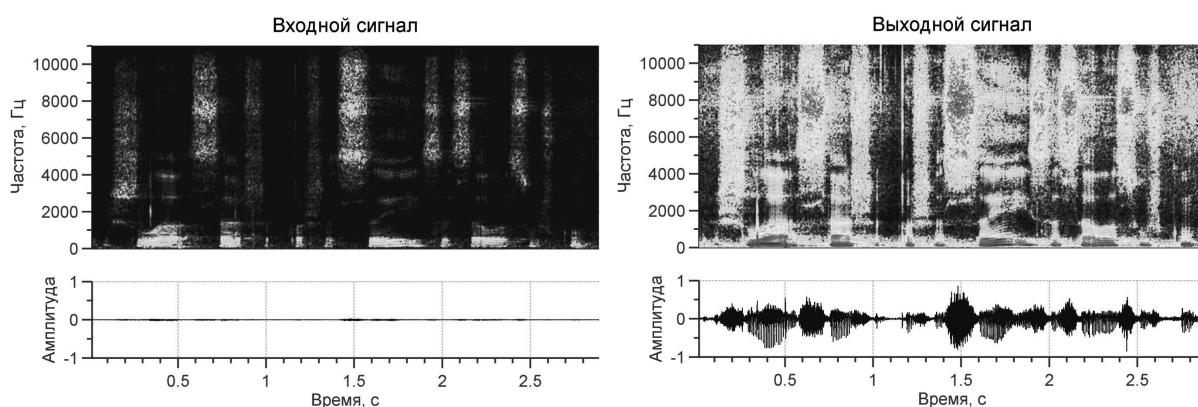


Рис. 6.12. Усиление речевого сигнала в слуховом аппарате

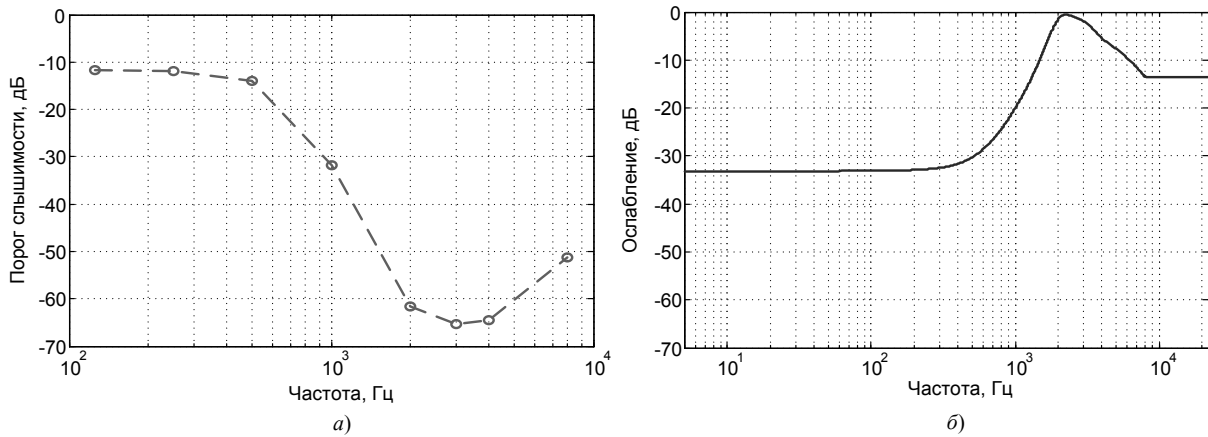


Рис. 6.13. Аудиометрия: а) аудиограмма пользователя; б) АЧХ фильтра-корректора

6.5.3. Пример обработки сигнала

Для иллюстрации работы предложенного алгоритма обработки сигнала в слуховом аппарате рассмотрим пример компенсации потери слуха. На рис. 6.13,а приведена аудиограмма пользователя (75 лет, 1-я степень тугоухости). По аудиограмме выполняется расчет фильтра-корректора, в котором заложены целевые коэффициенты усиления. АЧХ фильтра-корректора показана на рис. 6.13,б, для расчета которого использовалась формула Berger (таблица 6.2).

Чтобы качественно оценить результат работы предложенного алгоритма, выполнялось упрощенное моделирование потери слуха, в результате которого получался сигнал, воспринимаемый пациентом с ослабленным слухом. В общем случае система, моделирующая потерю слуха, должна отражать следующие явления, наблюдаемые при тугоухости [142]: 1) увеличение порога слышимости; 2) ускоренное нарастание громкости; 3) ограничение частотной избирательности; 4) компрессия в зависимости от уровня входного сигнала. В настоящей работе использовалась упрощенная система моделирования потери слуха, отражающая только явление увеличения порога слышимости. Для этого исходя из аудиограммы рассчитывался КИХ-фильтр методом частотной выборки. Вид речевого сигнала после прохождения через систему моделирования потери слуха показан на рис. 6.14,б. Соответствующий входной сигнал приведен на рис. 6.14,а.

Результат обработки речевого сигнала (рис. 6.14,а) разработанным алгоритмом коррекции слуха показан на рис. 6.15,а. Приведенные спек-

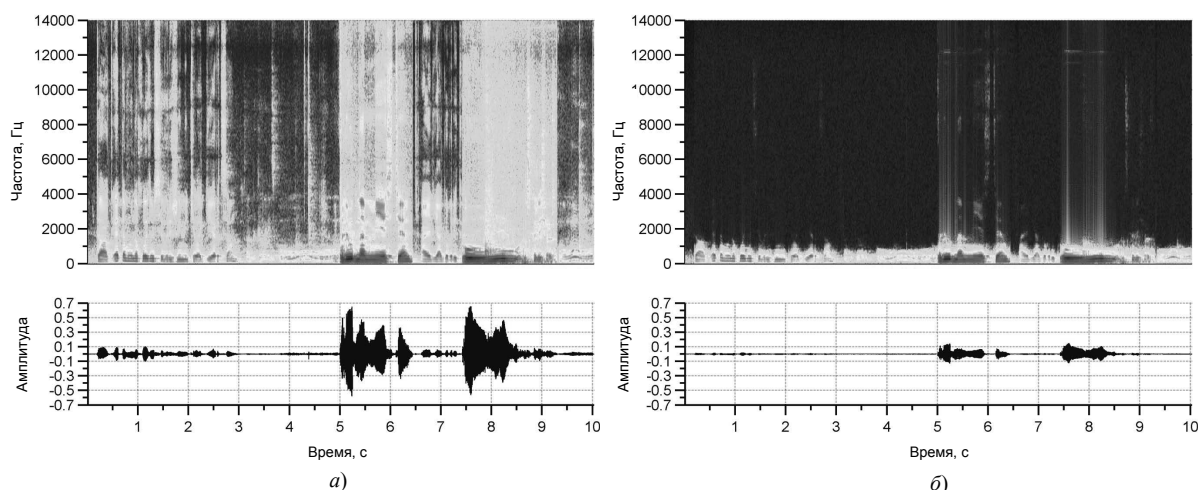


Рис. 6.14. Речевой сигнал: а) исходный; б) пропущенный через систему моделирования потери слуха

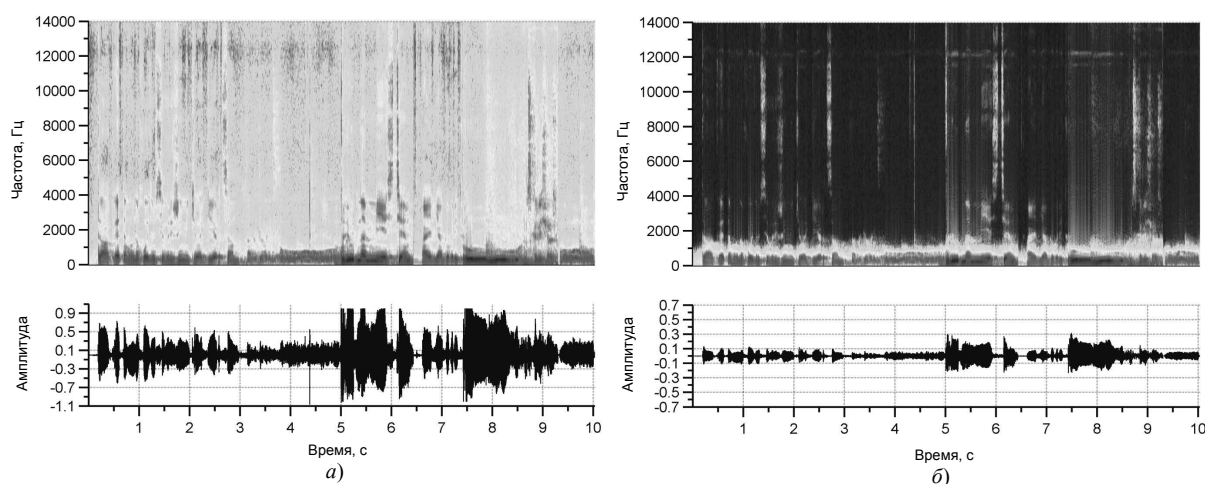


Рис. 6.15. Обработанный речевой сигнал: а) исходный; б) пропущенный через систему моделирования потери слуха

трограммы отражают существенное усиление высокочастотных составляющих речи. Вид сигнала во временной области показывает, что использование разработанного алгоритма компрессии позволяет сужать динамический диапазон сигнала, выполняя «выравнивание» громкости.

На рис. 6.15,б приводится обработанный сигнал, пропущенный через систему моделирования потери слуха. Сравнение данного сигнала с изображенным на рис. 6.14,б выявляет превосходство первого над вторым по субъективному уровню комфортности восприятия. Обработанный сигнал лучше воспринимается за счет суженного динамического диапазона (нет слишком тихих или слишком громких звуков), а также за счет частичного сохранения высокочастотных компонент речи, которые, как известно,

вносят большой вклад в разборчивость речи.

6.6. Резюме

В главе показана возможность применения мобильной вычислительной платформы iPhone для реализации цифрового слухового аппарата. Предложена схема обработки сигнала с малой алгоритмической задержкой, позволяющая выполнять коррекцию слуха при нейросенсорной тугоухости и учитывающая особенности вычислительной платформы. Подробно изложен алгоритм обработки, в основе которого лежат линейное частотно-зависимое усиление, а также широкополосная компрессия динамического диапазона сигнала. Приведены результаты обработки речевых сигналов с использованием предложенного алгоритма коррекции слуха.

ГЛАВА 7

АППАРАТНАЯ РЕАЛИЗАЦИЯ ВЫЧИСЛИТЕЛЬНЫХ ЭЛЕМЕНТОВ СЛУХОВОГО АППАРАТА НА ОСНОВЕ НЕРАВНОПОЛОСНОГО БАНКА ФИЛЬТРОВ

7.1. Структура процессора слухового аппарата

Учитывая алгоритм функционирования слухового аппарата на основе неравнополосного банка фильтров, разработанного в главе 5, предлагается следующая структура процессора слухового аппарата (рис. 7.1).

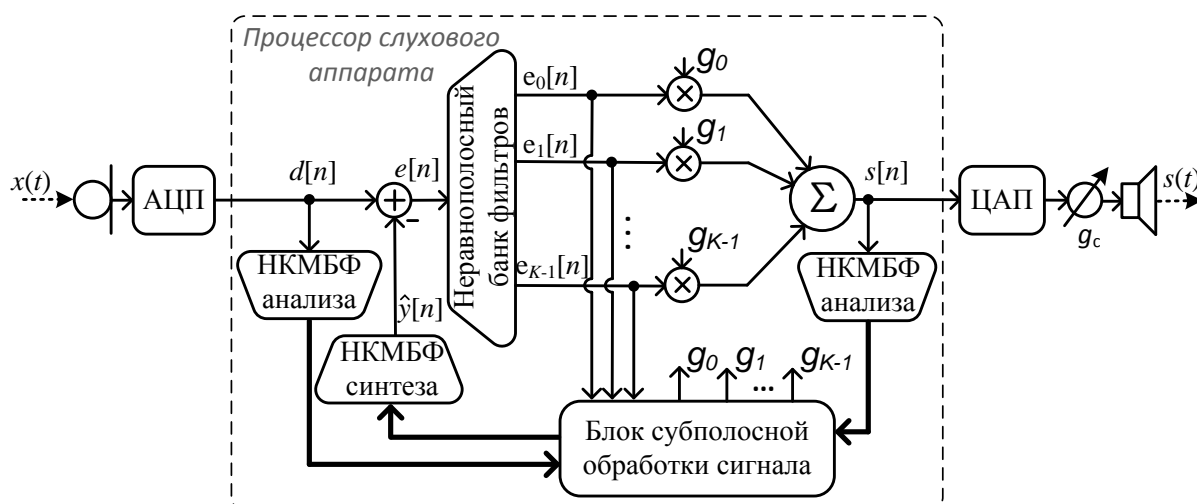


Рис. 7.1. Структурная схема слухового аппарата

Особенностью предлагаемой структуры является наличие в ней четырех неравнополосных банков фильтров, работающих параллельно, и вычислителя (блока субполосной обработки сигнала), задачей которого является выполнение следующих функций: обновление коэффициентов адаптивных фильтров, адаптивная фильтрация, выполнение алгоритмов шумоподавления, компрессии динамического диапазона (КДД) и компенсации потери слуха. В данной главе выполняется разработка способов реализации вычислительных элементов неравнополосных банков фильтров. Поскольку важнейшим требованием, предъявляемым к процессору слухового аппарата, является пониженное энергопотребление, то в качестве целевой вы-

Таблица 7.1. Вычислительная сложность слухового аппарата на основе неравнополосного банка фильтров: M – число каналов НКМБФ; N – порядок фильтра-прототипа НКМБФ; K – число каналов неравнополосного банка фильтров

Блок (алгоритм)	Вычислительная сложность
НКМБФ	$O(\frac{M}{2} \log_2 M + M + 2N)$
Неравнополосный банк фильтров	$O(\frac{K}{2} \log_2 K + 3M)$
Компенсация эффекта АОС	$O(\mathcal{O} \cdot (L + 2L))$
Шумоподавление + КДД	$O(4K + \frac{15}{R}K)$

числительной платформы выбираются программируемые логические микросхемы. Такой выбор позволяет реализовать необходимые в процессоре слухового аппарата неравнополосные банки фильтров в виде параллельно работающих структур. Основное внимание уделяется построению высокоуровневых вычислительных блоков, обеспечивающих гибкость и возможность быстрой адаптации к новым параметрам алгоритма и новым платформам. Под адаптацией понимается выбор числа каналов, порядка фильтра-прототипа, а также конфигурации полос банков фильтров в зависимости от частоты дискретизации входного сигнала.

Чтобы показать возможность эффективной аппаратной реализации предлагаемого процессора слуховых аппаратов, выполняется анализ вычислительной сложности. В табл. 7.1 приводятся вычислительная сложность основных блоков процессора слухового аппарата в терминах числа операций умножения на один входной отсчет. Полифазная реализация НКМБФ подразумевает использование быстрого алгоритма ДКП-4 ($O(\frac{M}{2} \log_2 M + M)$), умножение на коэффициенты фильтра-прототипа ($O(N)$) и выполнение фазового преобразования ($O(N)$). В зависимости от выбранных коэффициентов децимации/интерполяции НКМБФ субполосная система компенсации эффекта АОС требует $O(\mathcal{O}(L + 2L))$ операций умножения для реализации адаптивной фильтрации, где L – максимальная длина пути АОС в отсчетах, а \mathcal{O} – коэффициент передискретизации НКМБФ (см. выражение (2.30)).

Для оценки СПМ входного сигнала для алгоритмов шумоподавления

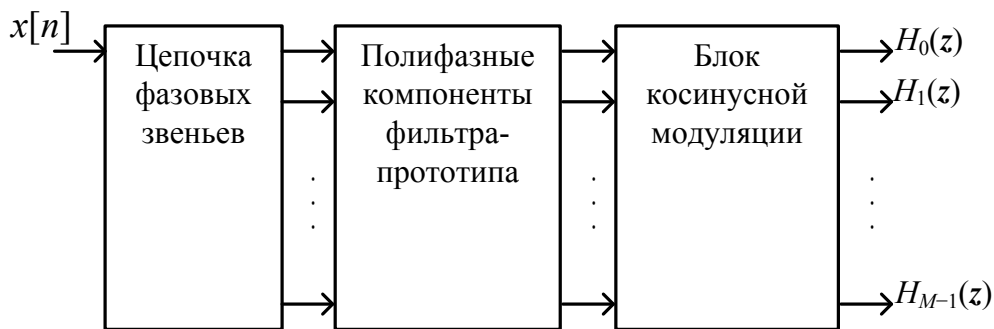


Рис. 7.2. Блоки НКМБФ

и КДД необходимо $O(4K)$ операций умножения. Перерасчет коэффициентов усиления g_0, \dots, g_{K-1} и реализация алгоритмов КДД и шумоподавления требует $O(15K)$ операций умножения. Однако это действие выполняется один раз для блока в R отсчетов. Таким образом, общая вычислительная сложность алгоритмов шумоподавления и КДД составляет $O(4K + \frac{15}{R}K)$. Из представленного описания следует, что в зависимости от значений параметров вычислительная сложность слухового аппарата растет линейно либо квазилинейно.

7.2. Реализация неравнополосного КМБФ

Рассмотрим реализацию неравнополосного КМБФ на примере реализации банка фильтров анализа (см. рис. 2.7). Структуру банка фильтров анализа можно разбить на три блока (рис. 7.2):

- цепочка фазовых звеньев;
- полифазные компоненты фильтра-прототипа;
- блок косинусной модуляции.

Входной отсчет, поступая в банк фильтров, сначала проходит по цепочке фазовых звеньев, затем выходные значения из цепочки поступают в полифазные компоненты фильтра-прототипа, после чего выходные сигналы формируются блоком косинусной модуляции. Далее рассматриваются различные варианты реализации каждого из трех блоков. В качестве примера выбран банк фильтров, используемый в прямом канале слухового аппарата, который имеет параметры: $M = 22$ — число полос; $N = 44$ — порядок фильтра-прототипа.

7.3. Цепочка фазовых звеньев

Как упоминалось ранее суть фазового преобразования заключается в замене всех элементов задержки цифровой системы на фазовые звенья (рис. 7.3).

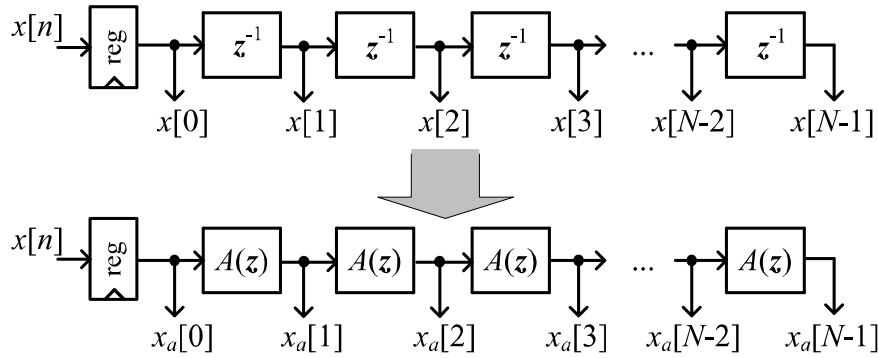


Рис. 7.3. Фазовое преобразование

Передаточная функция фазового звена задается как

$$A(z) = \frac{z^{-1} + \alpha}{1 + \alpha z^{-1}},$$

что во временной области соответствует разностному уравнению

$$y[n] = x[n - 1] + \alpha(x[n] - y[n - 1]). \quad (7.1)$$

Используя (7.1), видоизменим схему цепочки фазовых звеньев и представим её в виде комбинации умножителей, сумматоров и блоков задержки (рис. 7.4).

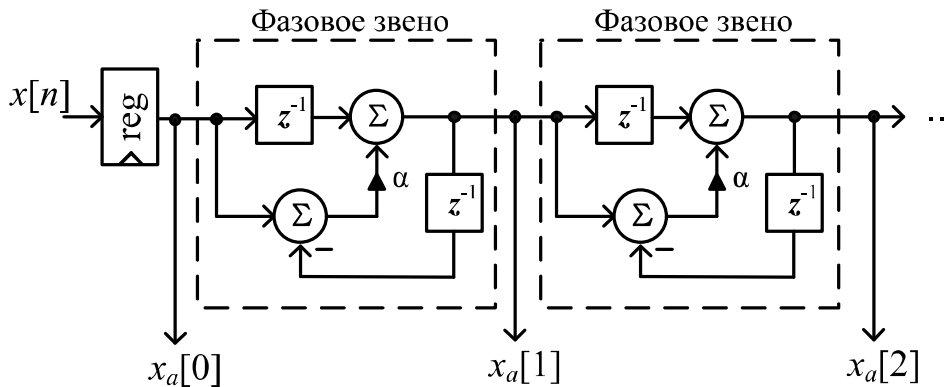


Рис. 7.4. Фазовое преобразование

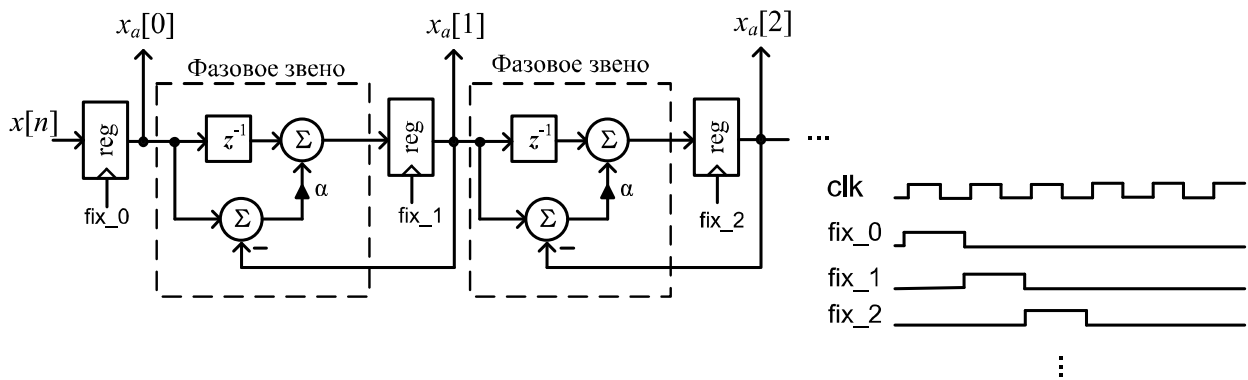


Рис. 7.5. Прямая реализация цепочки фазовых звеньев

Для НКМБФ необходимое количество фазовых звеньев определяется порядком фильтра-прототипа N . В рассматриваемом случае порядок фильтра-прототипа равен 44, это значит, что число фазовых звеньев будет равно 43. Рис. 7.4 показывает, что при поступлении нового отсчета $x[n]$ сигнал распространяется через все фазовые звенья. Если учесть, что задержка в одном фазовом звене составляет $2T_s + T_m$, где T_s — время суммирования (вычитания), а T_m — время умножения, то для всей цепочки мы получим время распространения сигнала равно $2(N - 1)T_s + (N - 1)T_m$. Такая задержка не допустима при аппаратной реализации процессора НКМБФ на ПЛИС. Простой способ, который можно предложить для устранения указанного недостатка, представлен на рис. 7.5.

Идея прямой реализации заключается в том, чтобы фиксировать результаты вычислений после каждого фазового звена в регистре, тогда за N тактов `clk` можно получить все нужные значения выходов цепочки фазовых звеньев. При такой реализации общее количество регистров не увеличивается, однако возникает необходимость в построении дополнительного блока для получения сигналов `fix`. Однако, сократив задержку распространения сигнала, аппаратные затраты не уменьшились. В нашем случае, когда порядок фильтра-прототипа равен 44, для реализации цепочки потребуется 87 регистров, 86 сумматоров и 43 умножителей, что конечно неприемлемо много для реализации на ПЛИС. Для иллюстрации того, что такой способ реализации является слишком затратным, написана VHDL-модель устройства на рис. 7.5 для случая, когда $\alpha = 0,5$ (т.е. операция умножения заменяется сдвигом вправо на один разряд). Ниже

приведен отчет программы-синтезатора после отображения цепочки фазовых звеньев на целевую ПЛИС. Из отчета видно, что прямая реализация цепочки фазовых звеньев занимает 20% от ресурсов всей целевой ПЛИС.

```

Target Device   : xc4vlx25
Target Package  : ff668
Target Speed    : -12
Mapper Version  : virtex4 -- $Revision: 1.34.32.1 $
    
```

Design Summary

Logic Utilization:

Number of Slice Flip Flops:	2,800 out of	21,504	13%
Number of 4 input LUTs:	2,876 out of	21,504	13%

Logic Distribution:

Number of occupied Slices:	2,191 out of	10,752	20%
Total Number of 4 input LUTs:	2,876 out of	21,504	13%

Для реализации цепочки фазовых звеньев может быть предложен и другой способ. Представим цепочку фазовых звеньев в следующем виде (рис. 7.6).

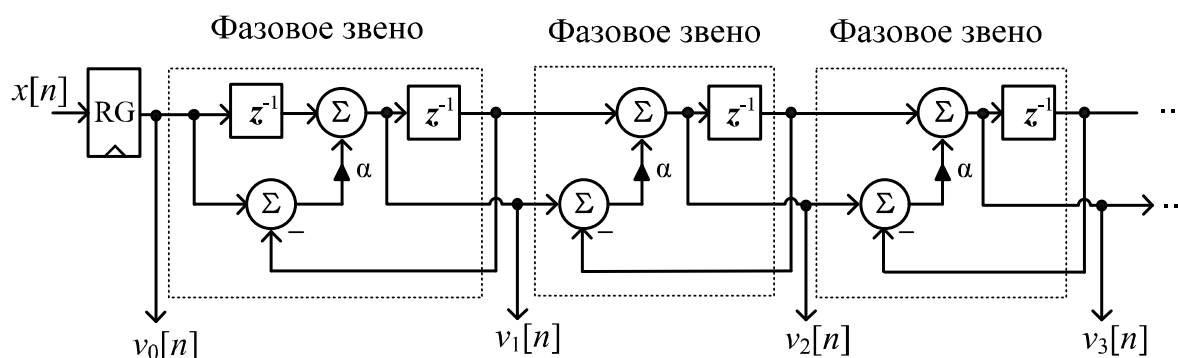


Рис. 7.6. Цепочка фазовых звеньев

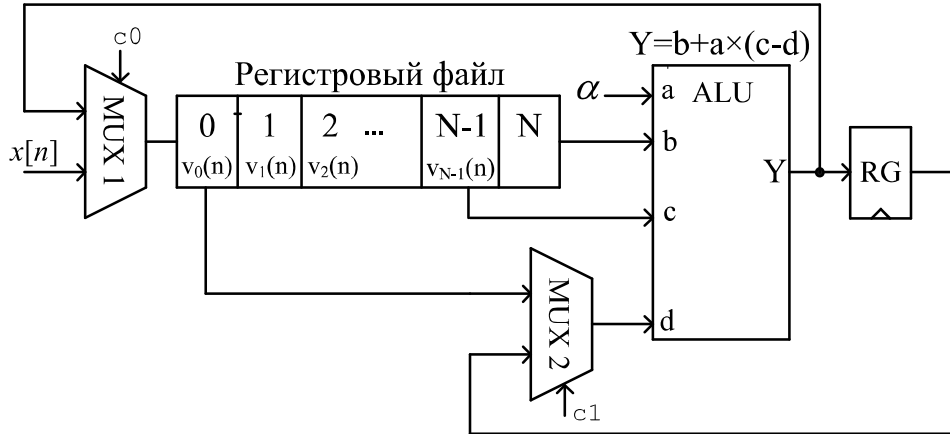


Рис. 7.7. Реализация цепочки фазовых звеньев

Для выходов $v_k[n]$ можно записать следующие выражения:

$$\begin{aligned}
 v_0[n] &= x[n]; \\
 v_1[n] &= v_0[n-1] + \alpha(v_0[n] - v_1[n-1]); \\
 v_2[n] &= v_1[n-1] + \alpha(v_1[n] - v_2[n-1]); \\
 &\dots \\
 v_{N-1}[n] &= F(\alpha, v_{N-2}[n-1], v_{N-2}[n], v_{N-1}[n-1]),
 \end{aligned} \tag{7.2}$$

где $F(a, b, c, d) = b + a(c - d)$.

На основании выражений (7.2) предлагается следующая реализация цепочки фазовых звеньев (рис. 7.7). Похожее устройство разрабатывалось в [69], однако в [69] структура цепочки фазовых звеньев была отличной от приведенной на рис. 7.6.

Поясним работу устройства на рис. 7.7. При поступлении нового входного отсчета $x[0]$ сигнал c_0 устанавливается в 1, т.о. входной отсчет попадает в регистровый файл. В следующем такте сигнал c_0 устанавливается в 0, а сигнал c_1 в 1. В этом же такте в ALU вычисляется выходной сигнал $v_1[0]$. Далее сигналы c_0, c_1 устанавливаются равными нулю, а устройство продолжает вычислять выходные сигналы цепочки фазовых звеньев. На N -ом такте работа устройства приостанавливается — в этот момент регистровый файл содержит все выходные сигналы цепочки фазовых звеньев.

После того как все выходные сигналы цепочки фазовых звеньев будут переданы в блок вычисления полифазных компонентов, работа устройства

возобновляется. Вычисления, производимые в ALU на $N + 1$ такте, ни на что не влияют, поскольку управляющий сигнал $c0$ в этот момент времени устанавливается в 1 и в регистровый файл попадает новый отсчет входного сигнала. А в следующем такте управляющий сигнал $c1$ устанавливается в 1, поэтому выходное значение ALU на $N + 1$ такте не влияет на работу устройства. Далее устройство работает по описанному выше алгоритму.

С другой стороны цепочку фазовых звеньев можно реализовать в виде вычислительной процедуры на универсальном процессоре [36]. Один из вариантов такой процедуры приведен ниже (предполагается, что R_1, R_2 и R_3 являются регистрами общего назначения процессора)

```

1:  $R_1 = x[n]$ 
2:  $R_2 = v[0]$ 
3:  $v[0] = R_1$ 
4: for  $i = 0, 1, \dots, N - 1$  do
5:    $R_3 = R_2$ 
6:    $R_2 = v[i]$ 
7:    $R_1 = (R_2 - R_1)\alpha + R_3$ 
8:    $v[i] = R_1$ 
9: end for

```

7.4. Блок полифазных компонент фильтра-прототипа

Следующим блоком после цепочки фазовых звеньев в процессоре НКМБФ является блок полифазных компонент фильтра-прототипа. Входом полифазных компонент служат выходные сигналы цепочки фазовых звеньев. Общая блок-схема формирования выхода i -го полифазного компонента показана на рис. 7.8.

Каждый полифазный компонент фильтра прототипа представляет собой КИХ-фильтр низкого порядка, который может быть эффективно реализован при помощи распределенной арифметики на памяти [143, 144]. На рис. 7.9 приведена структурная схема полифазного компонента, реализо-

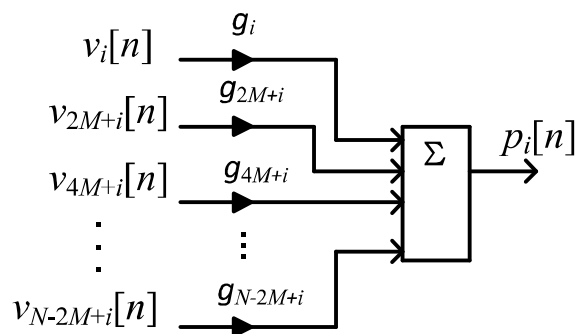


Рис. 7.8. Реализация цепочки фазовых звеньев

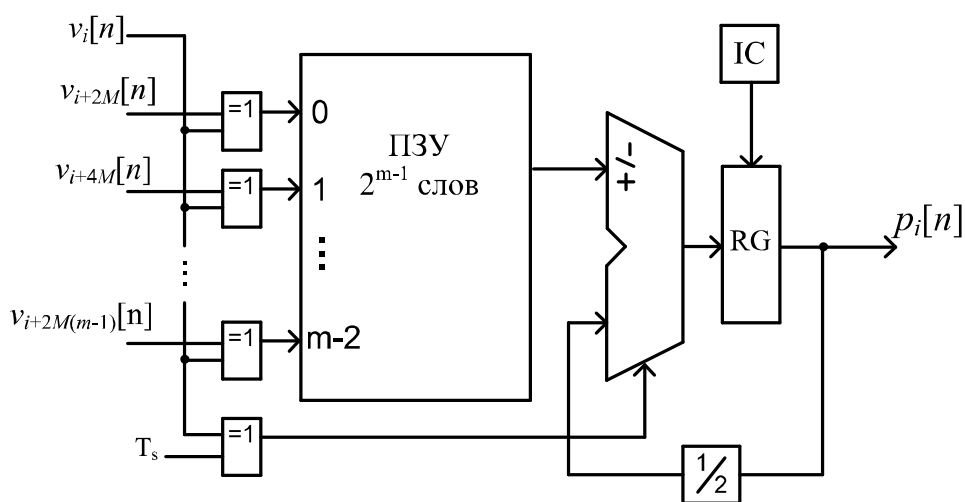


Рис. 7.9. Реализация цепочки фазовых звеньев

ванного при помощи распределенной арифметики [145].

На вход схемы бит за битом начиная с младшего разряда поступают данные из цепочки фазовых звеньев, и после шестнадцати тактов в регистре RG будет храниться результат. Устройство управления для этой схемы может быть выполнено на основе счетчика. В начале работы счетчик обнуляется, затем пробегает значения от 0 до 15, и во время последнего такта генерируется сигнал T_s .

Блок вычисления полифазных компонент так же, как и цепочка фазовых звеньев, может быть реализован в виде вычислительной процедуры на универсальном процессоре [36]. Псевдокод, реализующий блок полифазных компонент, приводится ниже.

В приведенном алгоритме $g[i + 2Mk]$ — коэффициенты фильтра-прототипа; $p_i[n]$ — выходной сигнала i -го полифазного компонента.

```
1: for  $i = 0, 1 \dots, 2M - 1$  do
2:    $R_1 = 0$ 
3:   for  $k = 0, 1 \dots, (N/2M) - 1$  do
4:      $R_2 = v[i + 2Mk]$ 
5:      $R_3 = g[i + 2Mk]$ 
6:      $R_1 = R_2 \times R_3 + R_1$ 
7:   end for
8:    $p_i[n] = R_1$ 
9: end for
```

7.5. Блок косинусной модуляции

После блока полифазных компонент в структуре НКМБФ следует блок косинусной модуляции. Частично вопрос реализации косинусной модуляции рассмотрен в подразделе 2.1.3., где показано, что главным элементом блока косинусной модуляции является быстрый алгоритм ДКП-4. Поэтому в дальнейшем подробно рассматривается вопрос получения схемы эффективной реализации ДКП [146].

Умножение на иррациональные множители является основной задачей при практической реализации быстрых алгоритмов ДКП. Этот вопрос важен, поскольку умножение чисел с плавающей запятой является довольно медленной операцией при аппаратной и программной реализациях. Структура многих быстрых алгоритмов ДКП содержит вычислительные блоки двумерного вращения плоскости (*plane rotation block*), в которых сосредоточены все иррациональные множители [95, 97]. Один из возможных эффективных подходов к реализации с фиксированной точкой операции вращения плоскости на основе лестничных структур предложен в [98]. Другой путь ухода от операции умножения с плавающей запятой является использование CORDIC алгоритмов [99]. Этот подход ведет к получению эффективных аппроксимаций 8-точечных ДКП с сокращенным критическим путем [100, 147]. Недостаток указанных подходов в используемом в них алгоритме Лоффлера, который позволяет получить быстрый алгоритм только, когда размера преобразования равен 4, 8 или 16 [98].

7.5.1. Быстрый алгоритм ДКП с использованием техники кодирования алгебраическими числами

Предлагается подход к реализации быстрых алгоритмов ДКП-2/ДКП-4, полученных в главе 4. Для получения аппаратных схем быстрых алгоритмов ДКП на основе арифметики с фиксированной запятой и обладающих высокой точностью предлагается использовать технику кодирования алгебраическими числами [148]. В отличие от [148] в настоящей работе выводится общая схема кодирования алгебраическими числами для полученных в главе 4 быстрых алгоритмов ДКП-2/ДКП-4.

В [148,149] предложена схема кодирования алгебраическими числами базисных функций ДКП для получения параллельных структур ДКП-процессоров с малой вычислительной сложностью. Основная идея техники кодирования алгебраическими числами состоит в том, чтобы представить все иррациональные множители в виде полиномов от z с целыми коэффициентами, где z — алгебраическое число.

Алгебраические числа. В общем случае алгебраическое число представляет собой корень полинома с коэффициентами из кольца целых чисел \mathbb{Z} . Например, пусть $z = 2 \cos(\pi/8) = \sqrt{2 + \sqrt{2}}$, тогда z является корнем полинома $p(z) = z^4 - 4z^2 + 2$. Если z присоединить к целым числам, то получающееся кольцо алгебраических чисел обозначается как $\mathbb{Z}[z]$. Кольцо $\mathbb{Z}[z]$ может рассматриваться как векторное пространство из полиномов от z , степень которых меньше 3, а коэффициенты являются целыми числами. Сложение и умножение элементов $\mathbb{Z}[z]$ производится по модулю полинома $p(z)$.

Кодирование алгебраическими числами. Для получения схемы кодирования алгебраическими числами для конкретного алгоритма необходимо выбрать подходящее алгебраическое число z и унитарный полином $p(z)$, такой что $p(z) = 0$. Алгебраическое число z должно выбираться таким, чтобы все иррациональные множители алгоритма могли быть выражены через полиномы от z с целыми коэффициентами. Рассмотрим схему кодирования алгебраическими числами для быстрого алгоритма ДКП-4,

полученного в подразделе 4.4.3.. Можно заметить, что все множители данного алгоритма имеют вид $2 \cos \frac{k\pi}{2n}$, где $0 \leq k < n$. Поэтому разумно принять $z = 2 \cos \frac{\pi}{2n}$. В этом случае полином $p(z)$ и все множители $2 \cos \frac{k\pi}{2n}$ выражаются при помощи полиномов Чебышева первого рода:

$$p(z) = 2T_n(z/2);$$

$$2 \cos \frac{k\pi}{2n} = 2T_k(z/2).$$

Необходимо отметить, что для данного способа полностью решается проблема погрешности представления, поскольку все иррациональные множители в данном случае представляются с абсолютной точностью.

Реконструкция результата. Используя технику кодирования алгебраическими числами, выходные результаты преобразования имеют полиномиальный вид:

$$f(z) = \sum_{i=0}^{n-1} a_i z^i.$$

Необходимо выполнить реконструкцию для того, чтобы перевести $f(z)$ в двоичный вид. Этот шаг может быть эффективно выполнен с использованием схемы Горнера [150]:

$$f(z) = (\dots (a_{n-1}z + a_{n-2})z + a_{n-3})z + \dots + a_1)z + a_0.$$

Для выполнения реконструкции необходимо получить аппроксимацию z в виде двоичного числа.

7.5.2. Реализация 16-точечного ДКП-2

Рассмотрим работу метода кодирования алгебраическими числами на примере быстрого алгоритма 16-точечного ДКП-2, граф-схема которого показана на рис. 4.4. Для реализации алгоритма получены схемы кодирования алгебраическими числами для 2-точечного, 4-точечного и 8-точечного ДКП-4, которые входят в состав быстрого алгоритма 16-точечного ДКП-2 (табл. 7.2).

Пример вычислений с использованием алгебраических чисел, при реализации 4-точечного ДКП-4 показан на рис. 7.10. Входные данные имеют

Таблица 7.2. Схемы кодирования алгебраическими числами для DCT-4₂, DCT-4₄ and DCT-4₈

Представление множителей для DCT-4₂ ($z = 2 \cos(\pi/4)$)

$$p(z) = z^2 - 2$$

$$2 \cos(\pi/4) = z$$

Представление множителей для DCT-4₄ ($z = 2 \cos(\pi/8)$)

$$p(z) = z^4 - 4z^2 + 2$$

$$2 \cos(\pi/8) = z$$

$$2 \cos(2\pi/8) = z^2 - 2$$

$$2 \cos(3\pi/8) = z^3 - 3z$$

Представление множителей для DCT-4₈ ($z = 2 \cos(\pi/16)$)

$$p(z) = z^8 - 8z^6 + 20z^4 - 16z^2 + 2$$

$$2 \cos(\pi/16) = z$$

$$2 \cos(2\pi/16) = z^2 - 2$$

$$2 \cos(3\pi/16) = z^3 - 3z$$

$$2 \cos(4\pi/16) = z^4 - 4z^2 + 2$$

$$2 \cos(5\pi/16) = z^5 - 5z^3 + 5z$$

$$2 \cos(6\pi/16) = z^6 - 6z^4 + 9z^2 - 2$$

$$2 \cos(7\pi/16) = z^7 - 7z^5 + 14z^3 - 7z$$

обыкновенный скалярный вид, а выходные представляют собой полиномы третьей степени. Для получения двоичных эквивалентов выходных значений к ним требуется применить операцию реконструкции.

При выполнении реконструкции выходных данных 2-точечного, 4-точечного и 8-точечного ДКП-4 использовались следующие аппроксимации

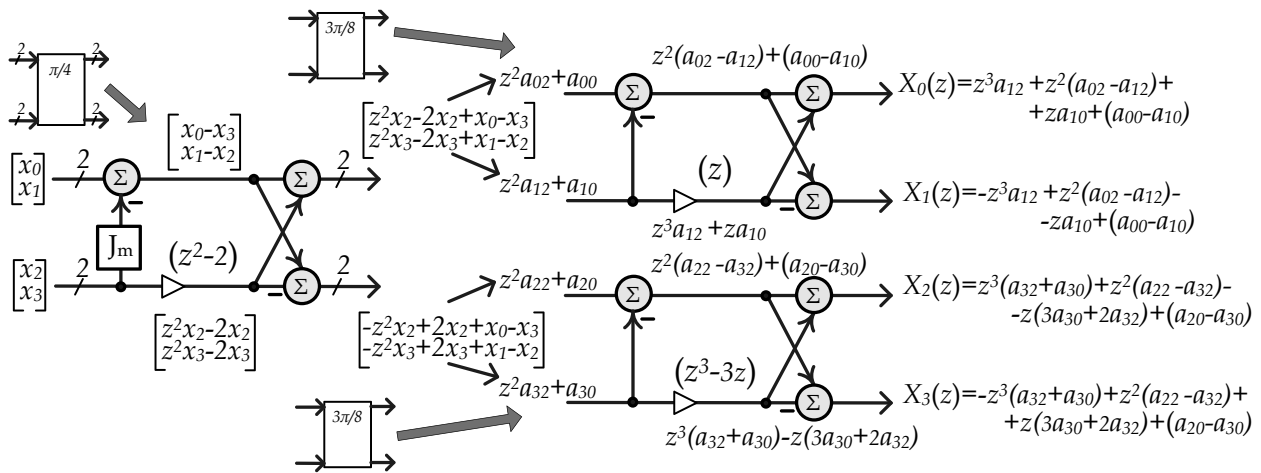


Рис. 7.10. Реализация 4-точечного ДКП-4 при помощи техники кодирования алгебраическими числами

алгебраического числа z :

$$\begin{aligned}
 2 \cos(\pi/4) &= 1 + 2^{-1} - 2^{-3} + 2^{-5} + 2^{-7}; \\
 2 \cos(\pi/8) &= 2 - 2^{-3} - 2^{-5} + 2^{-8}; \\
 2 \cos(\pi/16) &= 2 - 2^{-5} - 2^{-7} + 2^{-11}.
 \end{aligned}
 \tag{7.3}$$

На рис. 7.11 показана схема, выполняющая реконструкцию значений для 4-точечного ДКП-4. Коэффициенты полинома a_3, \dots, a_0 поступают на схему последовательно. Результат формируется после четырех циклов работы устройства. Необходимо отметить, что реконструкция вносит определенные ошибки округления, однако они появляются на самом последнем этапе преобразования, а не распространяются от узла к узлу быстрого алгоритма, как при обычных вычислениях с фиксированной запятой. Схемы реконструкции для алгоритмов 2-точечного и 8-точечного ДКП-4 имеют аналогичную структуру.

Оценка производительности реализации 16-точечного ДКП-2.

Для оценки производительности предложенной реализации 16-точечного ДКП-2, использующей алгебраические числа, воспользуется параметром эффективности кодирования (*coding gain*). Эффективность кодирования – важнейший параметр преобразования, используемый в приложениях сжатия информации [54]. Преобразование, для которого значение параметра эффективности кодирования выше, способно концентрировать больше

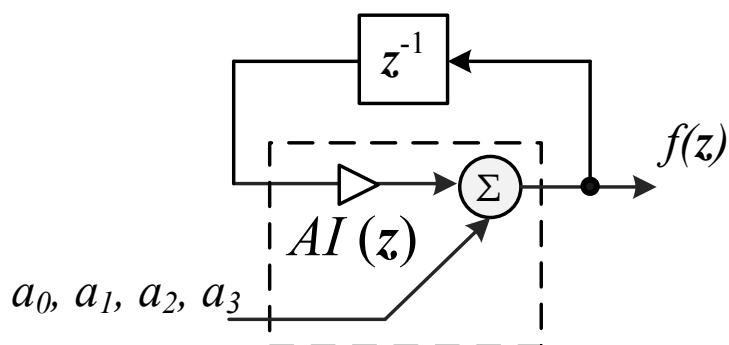


Рис. 7.11. Преобразования числа из полиномиальной формы в двоичную

энергии сигнала с меньшим числом коэффициентов. В результате этого при кодировании/декодировании изображений пиковое отношение сигнала/шум будет иметь большее значение. Параметр эффективности кодирования определяется как [54, 98]

$$C_g \triangleq 10 \log_{10} \frac{\sigma_x^2}{\left(\prod_{k=0}^{n-1} \sigma_{x_k}^2 \|f_k\|^2 \right)^{\frac{1}{n}}},$$

где n – число базисных функций преобразования;

σ_x^2 – дисперсия входного сигнала;

$\sigma_{x_k}^2$ – дисперсия k -го выходного сигнала;

$\|f_k\|^2$ – норма k -й базисной функции преобразования.

В качестве входного сигнала используется процесс Гаусса–Маркова с нулевым средним, единичной дисперсией и коэффициентом корреляции $\rho=0,95$. Известно [54], что оптимальным является базис Карунена–Лоэва, составляющий преобразование Карунена–Лоэва. Однако, априори базис Карунена–Лоэва недостаточно хорошо структурирован, поэтому на практике его часто стремятся аппроксимировать более структурированным базисом, который допускает более быстрые алгоритмы разложения. В табл. 7.3 приведены значения эффективности кодирования для 16-точечных преобразований Карунена–Лоэва, ДКП-2, binDCT [98] и предложенной реализации ДКП-2 с использованием алгебраических чисел (в таблице приведены два варианта алгоритма – при использовании 9 и 12 бит для представления дробной части данных).

При использовании 9 бит для представления дробной части данных в предложенной реализации ДКП-2 получено неожиданно высокое значение

Таблица 7.3. Значения эффективности кодирования для 16-точечных преобразований

Преобразование	Эффективность кодирования C_g , дБ
Карунена-Лоэва	9,4781
ДКП-2	9,4555
binDCT [98]	9,4499
Предложенная реализация ДКП (9 бит)	9,4567
Предложенная реализация ДКП (12 бит)	9,4553

эффективности кодирования. Этот эффект объясняется на рис. 7.12, где показаны АЧХ 12-й базисной функции ДКП-2 и предложенной реализации, использующей алгебраические числа. Четвертый боковой лепесток $H_{12}(e^{j\omega})$ имеет большее ослабление у предложенной реализации алгоритма, чем у оригинального ДКП-2. Таким образом, предложенная техника реализации быстрых алгоритмов ДКП позволяет получить численно-эффективные алгоритмы с высокими значениями показателя эффективности кодирования.

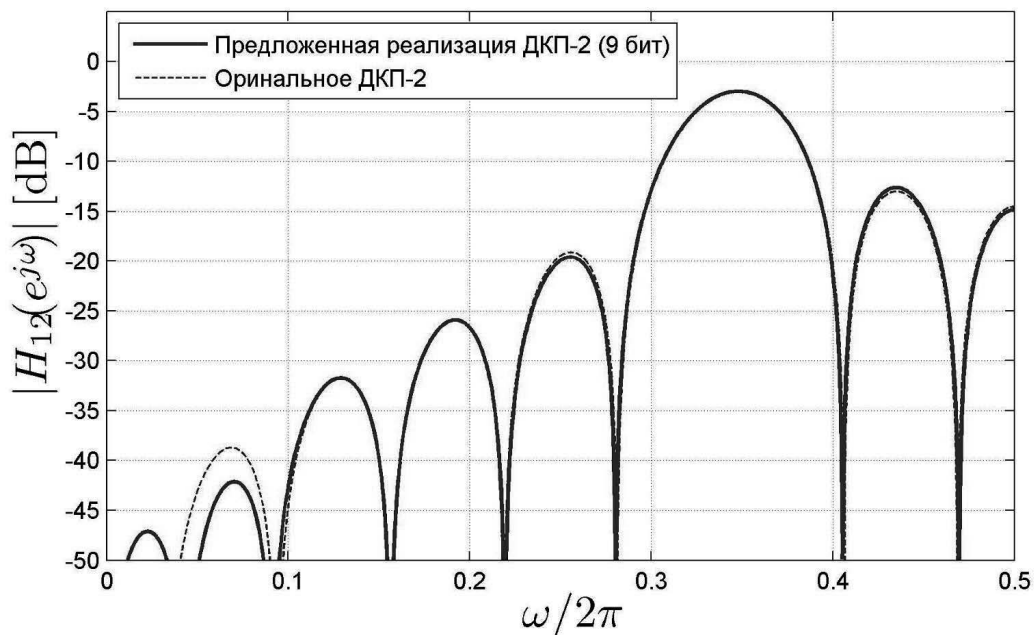


Рис. 7.12. АЧХ 12-й базисной функции ДКП-2

7.5.3. IP-ядра быстрого алгоритма ДКП

Для реализации быстрого алгоритма ДКП, использующего технику кодирования алгебраическими числами, разработаны и описаны на языке VHDL IP-ядра (*Intellectual Property*) этого алгоритма. В качестве примера рассмотрим быстрый алгоритм ДКП-4₈ (рис. 7.13), который необходим для построения НКМБФ в системе подавления акустической обратной связи. Для реализации базового блока быстрого алгоритма ДКП-4₈ выполнены несколько типов IP-ядер, отличающиеся по числу входов. Каждое IP-ядро характеризуется полиномиальным множителем и порядком полиномов, поступающих на его вход. В табл. 7.4 приведены выражения, по которым выполняется умножение в каждом базовом блоке быстрого алгоритма ДКП-4₈. В качестве примера рассмотрим работу базовых блоков, выполняющих умножение на полиномы, соответствующие множителям $2 \cos \frac{\pi}{4}$ и $2 \cos \frac{3\pi}{8}$.

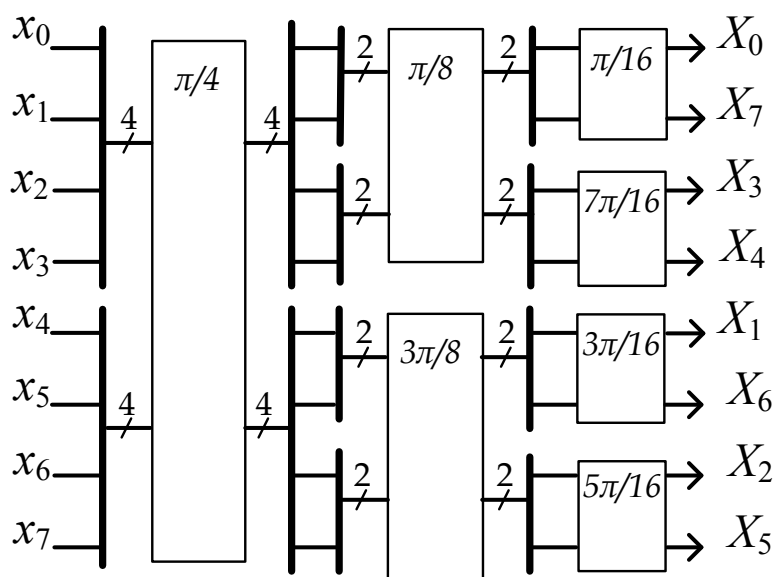


Рис. 7.13. Граф-схема быстрого алгоритма ДКП-4₈

На рис. 7.14 приведена схема реализации базового блока для угла $\pi/4$. Входом схемы служат скаляры x_0, x_1 и x_2 , а выходом являются полиномы $y_{0,0} + y_{0,2}z^2 + y_{0,4}z^4$ и $y_{1,0} + y_{1,2}z^2 + y_{1,4}z^4$, которые в схеме представлены только своими коэффициентами. В схеме кроме стандартных компонентов (сумматоров и вычитателей) имеется блок умножителя на константу $2 \cos \frac{\pi}{4}$. В табл. 7.4 приведены полином, соответствующий этому множителю, а

Таблица 7.4. Выражения для реализации умножителей в быстром алгоритме ДКП-4₈
 $(p(z) = z^8 - 8z^6 + 20z^4 - 16z^2 + 2)$

Множитель	Полином $f(z)$	Вход $a(z)$	Результат умножения $f(z)a(z) \bmod p(z)$
$2 \cos \frac{\pi}{16}$	z	$a_0 + a_2z^2 + a_4z^4 + a_6z^6$	$a_0z + a_2z^3 + a_4z^5 + a_6z^7$
$2 \cos \frac{\pi}{8}$	$z^2 - 2$	$a_0 + a_2z^2 + a_4z^4$	$2a_0 + (a_0 - 2a_2)z^2 + (a_2 - 2a_4)z^4 + a_4z^6$
$2 \cos \frac{3\pi}{16}$	$z^3 - 3z$	$a_0 + a_2z^2 + a_4z^4 + a_6z^6$	$(-2a_0 - 2a_2 - 4a_4) + (9a_0 + 14a_2 + 30a_4)z^2 - (6a_0 + 11a_2 + 26a_4)z^4 + (a_0 + 2a_2 + 5a_4)z^6$
$2 \cos \frac{\pi}{4}$	$z^4 - 4z^2 + 2$	a_0	$2a_0 - 4a_0z^2 + a_0z^4$
$2 \cos \frac{5\pi}{16}$	$z^5 - 5z^3 + 5z$	$(a_0 + a_2z^2 + a_4z^4 + a_6z^6)$	$(5a_0 - 2a_4 - 6a_6)z - (5a_0 - 5a_2 - 16a_4 - 46a_6)z^3 + (a_0 - 5a_2 - 15a_4 - 44a_6)z^5 + (a_2 + 3a_4 + 9a_6)z^7$
$2 \cos \frac{3\pi}{8}$	$z^6 - 6z^4 + 9z^2 - 2$	$a_0 + a_2z^2 + a_4z^4$	$(-2a_0 - 2a_2 - 4a_4) + (9a_0 + 14a_2 + 30a_4)z^2 - (6a_0 + 11a_2 + 26a_4)z^4 + (a_0 + 2a_2 + 5a_4)z^6$
$2 \cos \frac{7\pi}{16}$	$z^7 - 7z^5 + 14z^3 - 7z$	$a_0 + a_2z^2 + a_4z^4 + a_6z^6$	$(-7a_0 - 2a_2 - 2a_4 - 4a_6)z + (14a_0 + 9a_2 + 14a_4 + 30a_6)z^3 - (7a_0 + 6a_2 + 11a_4 + 26a_6)z^5 + (a_0 + a_2 + a_4 + 5a_6)z^7$

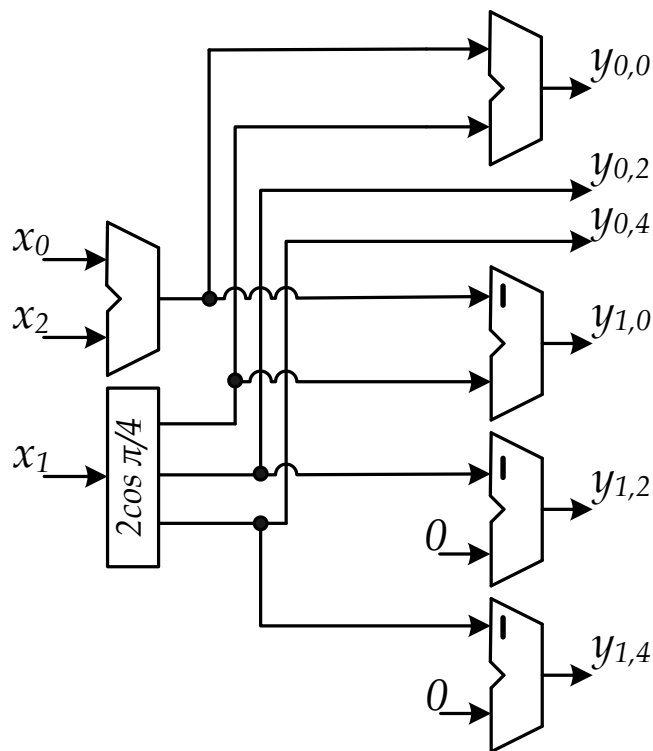


Рис. 7.14. Базовый блок быстрого алгоритма ДКП-4₈ ($\alpha = \frac{\pi}{4}$)

также математические действия, выполняемые внутри блока (последний столбец). Заметим, что умножение на фиксированные коэффициенты могут быть легко реализованы с использованием операций сдвига и сложения.

Рис. 7.15 изображает схему реализации базового блока для угла $3\pi/8$. В отличие от предыдущего блока, входом схемы служат полиномы $x_{0,0} + x_{0,2}z^2 + x_{0,4}z^4$ и $y_{1,0} + y_{1,2}z^2 + y_{1,4}z^4$, представленные своими коэффициентами. Выходом являются полиномы более высокой степени $y_{0,0} + y_{0,2}z^2 + y_{0,4}z^4 + y_{0,6}z^6$ и $y_{1,0} + y_{1,2}z^2 + y_{1,4}z^4 + y_{1,6}z^6$. Блок умножителя на константу $2 \cos \frac{3\pi}{8}$ выполнен в соответствии с табл. 7.4. Остальные базовые блоки реализованы аналогичным образом.

Как указывалось ранее, на выходе схемы быстрого алгоритма данные представлены в полиномиальной форме. Для перевода их к числовому виду необходим блок реконструкции результата FRS (*Final Reconstruction Step*). Для этой цели разработано устройство, схема которого показана на рис. 7.16. Преобразование результата из полиномиальной формы в числовую происходит последовательно во времени. В начале работы устройства FRS регистр RG сбрасывается при помощи сигнала *rst*. Далее, используя

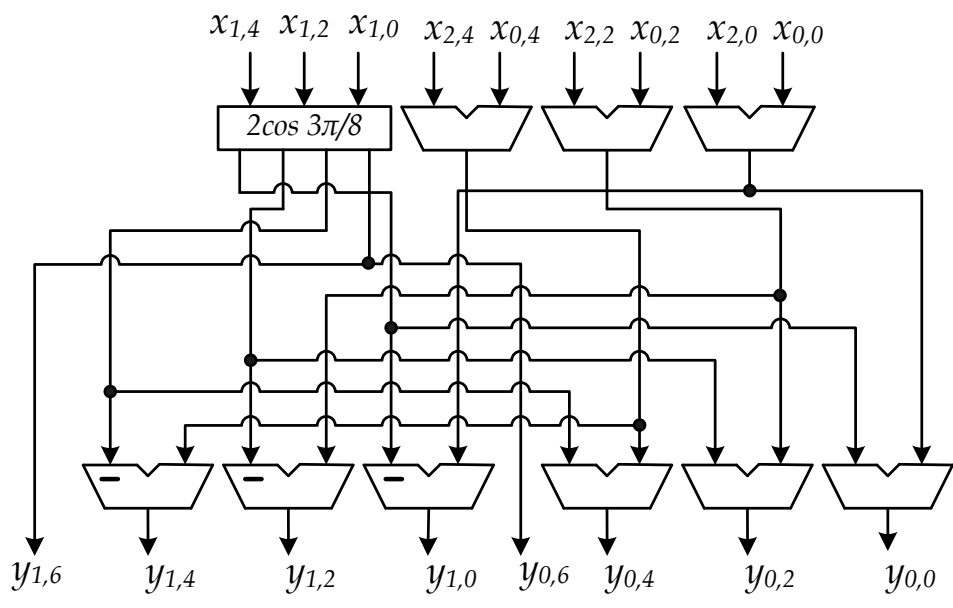


Рис. 7.15. Базовый блок быстрого алгоритма ДКП-4₈ ($\alpha = 3\pi/8$)

адресный сигнал мультиплексора *addr*, выбирается старший коэффициент полинома a_7 , который складывается с содержимым регистра RG, умноженным на алгебраическое число z . Затем выбирается следующий коэффициент a_6 и действия повторяются и т.д. После восьми тактов работы регистр RG содержит результат реконструкции. Для реализации блока, выполняющего умножение на z , использовалась аппроксимация (7.3).

Общая схема реализации быстрого алгоритма ДКП-4₈, использующая разработанные IP-ядра, приведена на рис. 7.17. Входами и выходами схемы являются скалярные величины, однако промежуточные данные представлены в полиномиальном виде (на схеме обозначены жирным шрифтом). В табл. 7.5 приведен общий вид полиномов на каждой ступени быстрого алгоритма. Первые три ступени алгоритма представляют собой сложную комбинационную схему без элементов памяти. Задержку вносят только блоки FRS, которые требуют дополнительных восемь тактов для формирования результата.

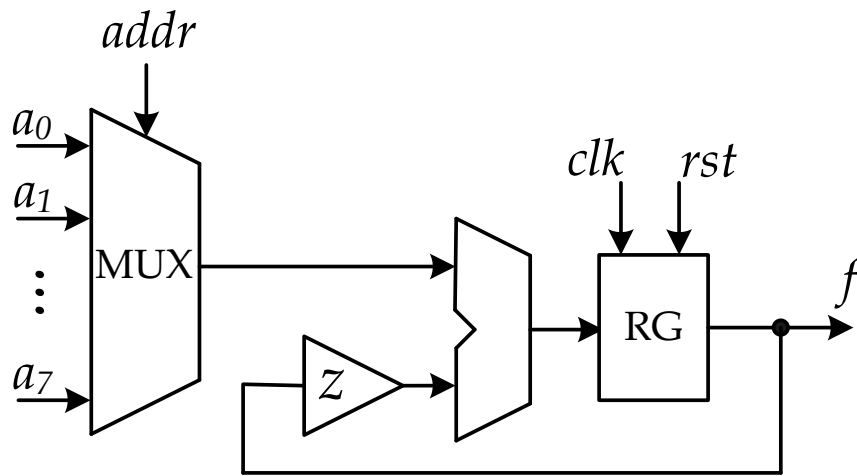


Рис. 7.16. Реализация блока реконструкции результата $a_7z^7 + a_6z^6 + a_5z^5 + a_4z^4 + a_3z^3 + a_2z^2 + a_1z + a_0$, $z = 2 \cos \frac{\pi}{16}$

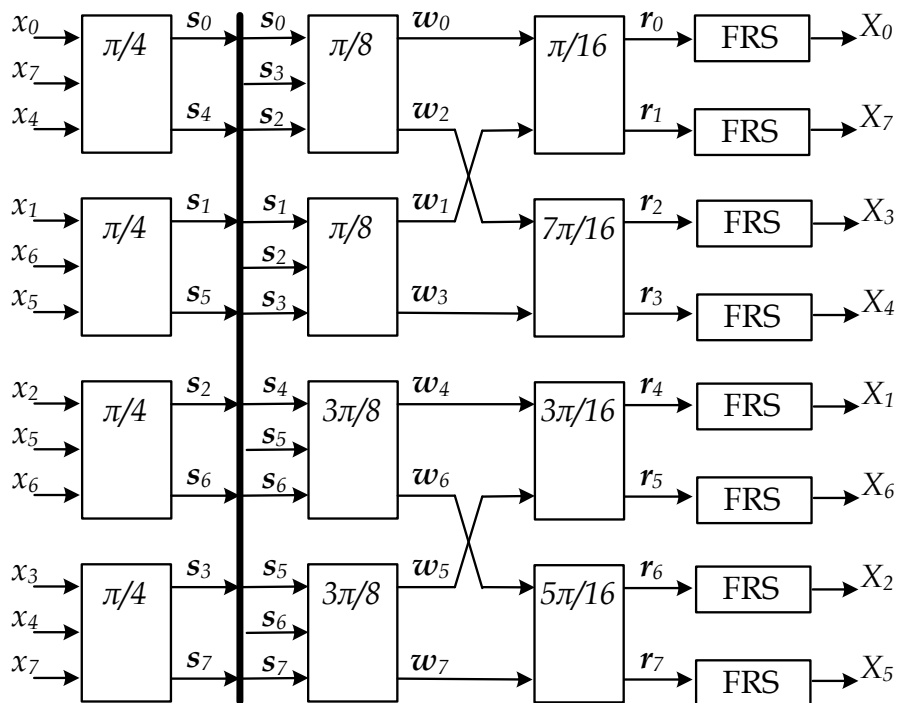


Рис. 7.17. Реализация быстрого алгоритма ДКП-4₈, использующая разработанные IP-ядра

Таблица 7.5. Представление данных в схеме быстрого алгоритма ДКП-8₄

Номер ступени	Вид полинома
1	$s_i = s_{i,0} + s_{i,2}z^2 + s_{i,4}z^4$
2	$w_i = w_{i,0} + w_{i,2}z^2 + w_{i,4}z^4 + w_{i,6}z^6$
3	$r_i = r_{i,0} + r_{i,1}z + r_{i,2}z^2 + r_{i,3}z^3 + r_{i,4}z^4 + r_{i,5}z^5 + r_{i,6}z^6 + r_{i,7}z^7$

7.6. Система быстрого прототипирования неравнополосного косинусно-модулированного банка фильтров

Разработанные в работе методы служат основой для создания среды быстрого прототипирования неравнополосных банков фильтров на процессоры слуховых аппаратов. Аппаратной основой для реализации процессора слухового аппарата выбирается ПЛИС. Такой выбор позволяет реализовать необходимые в процессоре слухового аппарата неравнополосные банки фильтров в виде параллельно работающих структур, что существенно сокращает тактовую частоту работы устройства и, как следствие, энергопотребление. Процессор слухового аппарата представляет собой четыре банка фильтров, работающих параллельно, а также вычислитель, задачей которого является расчет субполосных коэффициентов усиления, а также адаптация коэффициентов для системы компенсации АОС (см. рис. 7.1).

Для осуществления возможности сквозного проектирования разработаны параметризованные VHDL-модели всех блоков неравнополосного косинусно-модулированного банка фильтров, а именно: цепочки фазовых звеньев, полифазных компонент фильтра прототипа и быстрого алгоритма ДКП. Полученные VHDL-модели могут использоваться в системе автоматизированного проектирования неравнополосного банка фильтров. Общая схема процессора неравнополосного банка фильтров для слухового аппарата, получающаяся в результате объединения разработанных компонентов, показана на рис. 7.18. Входной сигнал поступает в цепочку фазовых звеньев, затем проходит через полифазные компоненты фильтра-прототипа и

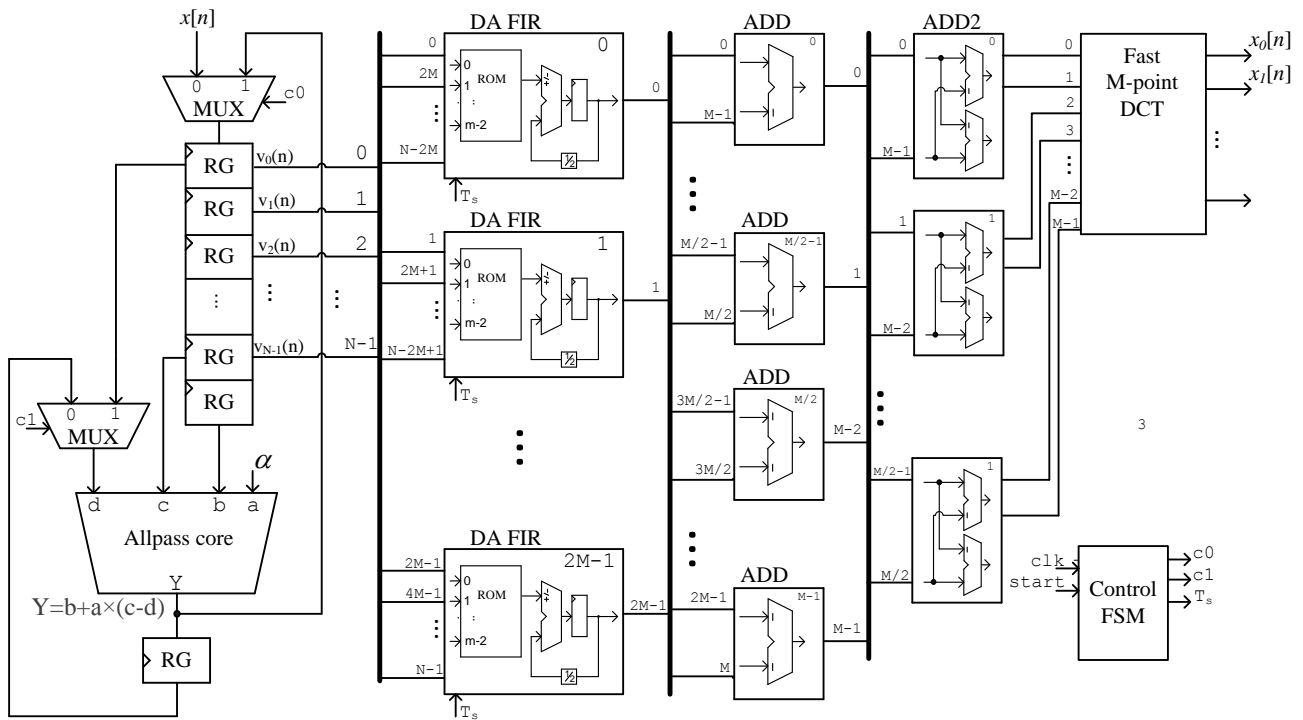


Рис. 7.18. Процессор неравнополосного косинусно-модулированного банка фильтров

поступает в блок суммирования, после чего результат формируется блоком быстрого алгоритма ДКП.

На рис. 7.19 суммирована последовательность действий, выполняемая в среде быстрого прототипирования для получения процессора неравнополосного банка фильтров. На начальном этапе задаются число каналов, порядок фильтра-прототипа и коэффициент деформации частотной оси. Все остальные шаги выполняются автоматически. Исходя из конфигурации рассчитываются коэффициенты децимации/интерполяции. Затем выполняется оптимизация фильтра-прототипа. После чего генерируется быстрый алгоритм ДКП. Затем происходит отображение алгоритмов всех блоков процессора банка фильтров, используя параметризованные VHDL-описания IP-ядер. Результатом всего процесса является синтезируемое описание процессора неравнополосного косинусно-модулированного банка фильтров. Полученный проект может быть верифицирован в системе ModelSim, а также синтезирован с использованием специальных САПР Xilinx ISE или Leonardo Spectrum. На рис. 7.20 показано рабочее место, оборудованное для выполнения быстрого прототипирования процессора неравнополосного косинусно-модулированного банка фильтров.

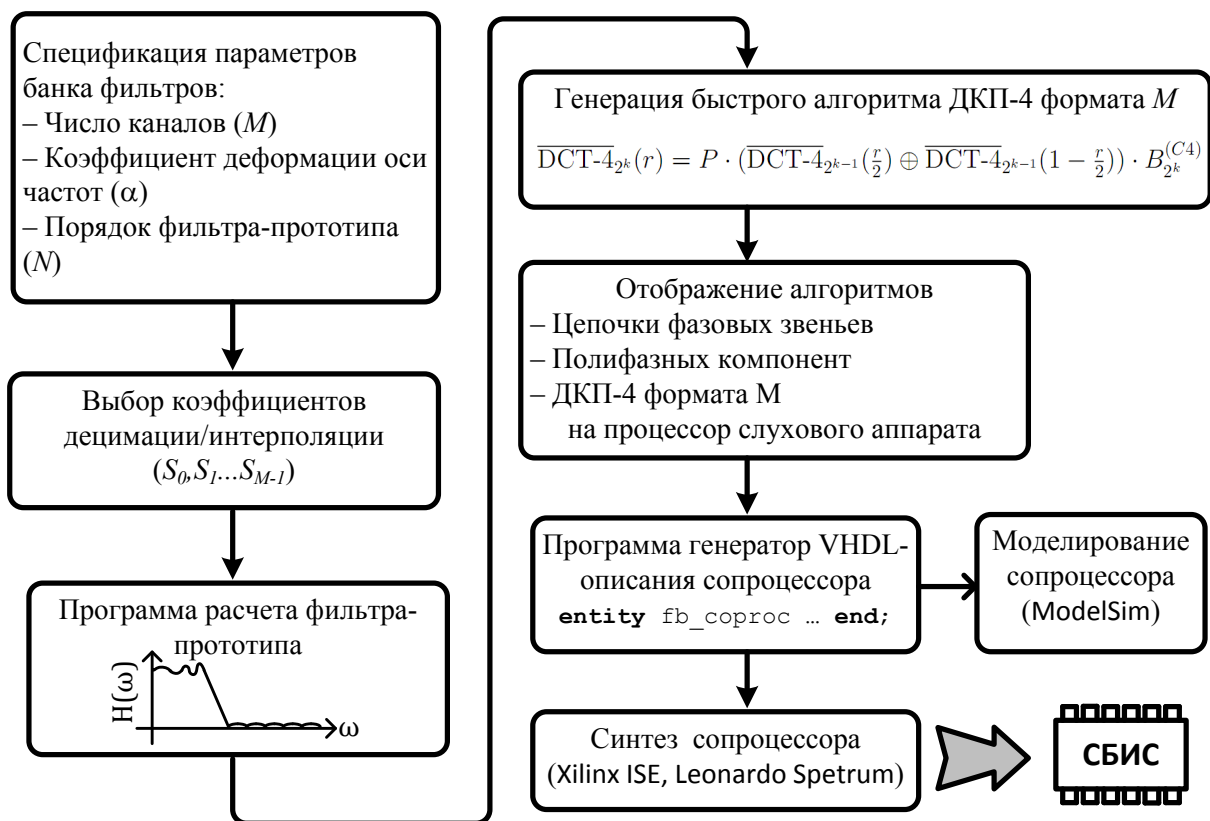


Рис. 7.19. Последовательность действий в системе быстрого прототипирования процессора неравнополосного косинусно-модулированного банка фильтров

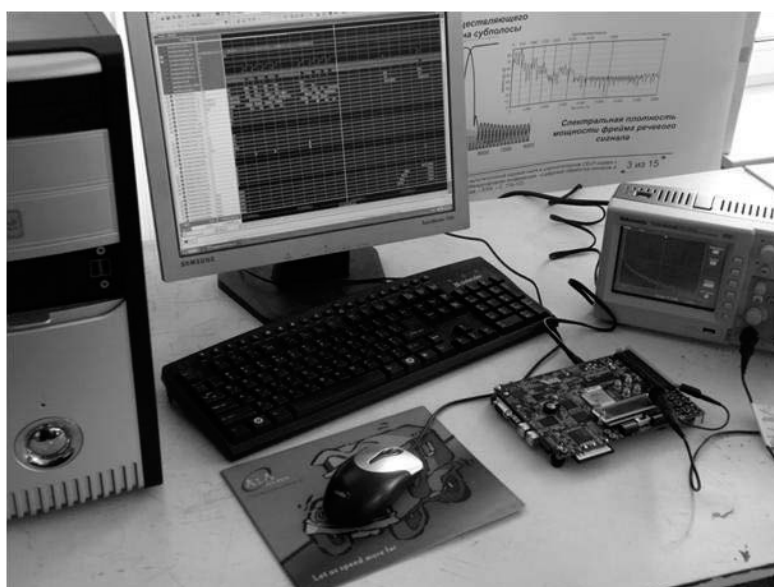


Рис. 7.20. Инструментальные средства для быстрого прототипирования неравнополосного банка фильтров

В качестве примера выполнено быстрое прототипирование процессора НКМБФ, используемого в субполосной системе подавления АОС. Данный

банк фильтров имеет восемь полос ($M = 8$), коэффициент деформации частотной оси $\alpha = 0,5354$ и порядок фильтра прототипа $N = 48$. В табл. 7.6 приведены аппаратные затраты ПЛИС (типа FPGA) на реализацию данного банка фильтров. В скобках для каждого вида ресурса приведен процент от общего имеющихся на кристалле.

Таблица 7.6. Ресурсы FPGA XC4VLX25 на реализацию процессора НКМБФ

Устройство	Число триггеров	Число LUT	Число секций
Цепочка ФЗ	808 (3%)	856 (3%)	429 (3%)
Полифазные фильтры	256 (1%)	303 (1%)	155 (1%)
Быстрый алгоритм ДКП-4 ₈	134 (1%)	3556 (16%)	1894 (17%)
Итого	1199 (5%)	4833 (22%)	2551 (23%)

На рис. 7.21 показана диаграмма затрат аппаратных ресурсов FPGA XC4VLX25 (количество секции Slice) на реализацию отдельных блоков процессора НКМБФ.



Рис. 7.21. Аппаратные ресурсы FPGA XC4VLX25 на реализацию процессора НКМБФ

7.7. Основные результаты и выводы

В главе рассмотрены вопросы практической реализации вычислительных элементов процессора слухового аппарата на основе неравнополосного косинусно-модулированного банка фильтров с субполосной компенсацией эффекта АОС.

Для реализации цепочки фазовых звеньев предложена структура устройства на основе регистрового файла, двух мультиплексоров и простейшего АЛУ [36]. Особенность реализации заключается в сокращении числа вычислительных элементов в N раз по сравнению с прямой реализацией (N – число фазовых звеньев в цепочке).

Предложена схема реализации полифазных компонент фильтра-прототипа НКМБФ на базе распределенной арифметики [145], которая позволяет учесть особенности аппаратной платформы ПЛИС и избежать использования умножителей.

Рассмотрены вопросы реализации быстрых алгоритмов ДКП-2/ДКП-4 с использованием алгебраических чисел. Предложена общая схема кодирования алгебраическими числами иррациональных множителей быстрых алгоритмов ДКП. Разработанный способ вычисления быстрых алгоритмов ДКП позволяют реализовать их эффективно на выбранной аппаратной платформе. Показано, что получающиеся в результате реализации схемы ДКП обладают точностью, которая превосходит существующие аналоги, что доказывается высокими значениями эффективности кодирования [146].

ЗАКЛЮЧЕНИЕ

К основным научным результатам, представленным в монографии можно отнести:

1. Метод расчета передискретизированного неравнополосного косинусно-модулированного банка фильтров [47], позволяющий уменьшить искажения, вносимые при децимации/интерполяции сигналов в каналах банка фильтров. Отличительной чертой метода является возможность учитывать при расчете конкретные значения коэффициентов децимации/интерполяции канальных сигналов, что позволяет обеспечить уровень искажений в банке фильтров порядка 0,005 дБ [75, 77].

2. Алгебраический метод синтеза быстрых алгоритмов дискретного косинусного преобразования (ДКП), особенностью которого является отсутствие ограничения на размер преобразования [89, 90].

3. Разработка обобщенной схемы кодирования алгебраическими числами иррациональных множителей быстрых алгоритмов ДКП [89] для исключения операций умножения, что позволяет эффективно реализовывать быстрые алгоритмы ДКП на ПЛИС. Показано, что точность генерируемых схем ДКП превосходит существующие аналоги. Это доказывается высокими значениями коэффициента эффективности кодирования (9,4567 дБ для 16-точечного ДКП) [146].

4. Способ субполосной компенсации эффекта АОС на основе неравнополосного косинусно-модулированного банка фильтров [74], не вносящий дополнительной задержки в прямой путь распространения сигнала в слуховом аппарате. Использование разработанного способа позволяет увеличить коэффициент усиления сигнала в слуховом аппарате на 3 дБ [51].

5. Структура слухового аппарата на основе неравнополосного банка фильтров с субполосной компенсацией эффекта акустической обратной связи, которая в отличие от существующих позволяет выполнить обработку сигнала с малой алгоритмической задержкой, около 4 мс [46, 51].

6. Разработаны параллельные структуры вычислительных блоков неравнополосного косинусно-модулированного банка фильтров [90, 145].

Полученные методы и алгоритмы применимы в различных областях цифровой обработки сигналов. В первую очередь это относится к методу

синтеза быстрых алгоритмов ДКП, поскольку это преобразование имеет широкое распространение и используется в таких областях, как сжатие изображений/видео, распознавание образов, запись медицинских сигналов (ЭЭГ, ЭКГ) и т.д. Разработанный НКМБФ может использоваться для построения психоакустических моделей, которые широко применяются при создании систем кодирования звука, конверсии голоса, в задачах шумоподавления.

ЛИТЕРАТУРА

1. Глухота и нарушения слуха. Информационный бюллетень № 300. [Электронный ресурс]. – 2012. – Режим доступа : <http://www.who.int/mediacentre/factsheets/fs300/ru/index.html>. – Дата доступа: 29.07.2012.
2. Лисовский В.А. Слуховые приборы и аппараты / В.А. Лисовский, В.А. Елисеев — М.: Радио и связь, 1991. — 192 с.
3. Kates J.M. Multichannel dynamic-range compression using digital frequency warping / J.M. Kates, K.H. Arehart // EURASIP Journal on Applied Signal Processing. — 2005. — no. 18. — P. 3003–3014.
4. Löllmann H.W. A warped low delay filter for speech enhancement / H.W. Löllmann, P. Vary // Proceedings of the International Workshop on Acoustic Echo and Noise Control. — Paris, France: 2006. — P. 1–4.
5. Bäuml R.W. Uniform polyphase filter banks for use in hearing aids: design and constraints / R.W. Bäuml, W. Sörgel // Proceedings of European Signal Processing Conference (EUSIPCO'2008). — Lausanne, Switzerland: 2008.
6. Pandey A. Low-delay signal processing for digital hearing aids / A. Pandey, V.J. Mathews // IEEE Transactions on Audio, Speech, and Language processing. — 2011. — Vol. 19, no. 4. — P. 699–710.
7. WOLA filter bank design requirements in hearing aids / M.R. Zurera, R.G. Pita, E.A. Cortizo et al. // Proceedings of International Conference on Pattern Recognition and Information Processing. — Minsk, Belarus: 2009. — P. 215–218.
8. Солонина А.И. Алгоритмы и процессоры цифровой обработки сигналов / А.И. Солонина, Д.А. Улахович, Л.А. Яковлев— СПб.: БХВ-Петербург, 2002. — 464 с.
9. Айфичер Э. Цифровая обработка сигналов: практический подход. Второе издание / Э. Айфичер, Б. Джервис— М.: Вильямс, 2004. — 992 с.
10. Фонлантен А. Слуховые аппараты / А. Фонлантен, Х. Арндт— Ростов н/Д.: Феникс, 2009. — 304 с.
11. Лабутин В.К. Слух и анализ сигналов / В.К. Лабутин, А.П. Молчанов— М.: Энергия, 1967. — 80 с.

12. Эфрусси М.М. Слуховые приборы и аудиометры / М.М. Эфрусси— М.: Энергия, 1975. — 96 с.
13. Коломийченко А.М. Атлас тональных аудиометрических исследований / А.М. Коломийченко, Н.С. Шейнман— Киев : Государственное медицинское издательство УССР, 1962. — 293 с.
14. Lippman R.P. Study of multichannel amplitude compression and linear amplification for persons with sensorineural hearing loss / R.P. Lippman, L.D. Braida, N.I. Durlach // Journal of the Acoustical Society of America. — 1981. — no. 69. — P. 524–534.
15. Dillon H. Hearing Aids / H. Dillon. — Boomerang Press, 2001. — 504 p.
16. Zwicker E. Analytical expressions for critical-band rate and critical bandwidth as a function of frequency / E. Zwicker, E. Terhardt // Journal of the Acoustical Society of America. — 1980. — Vol. 68. — P. 1523–1525.
17. Levitt H. Noise reduction in hearing aids: a review / H. Levitt // Journal of Rehabilitation Research and Development. — 2001. — Vol. 38. — P. 111–121.
18. Loizou P. Speech enhancement: theory and practice / P. Loizou. — Boca Raton, FL: CRC Press, 2007. — 632 p.
19. Moore B. Speech processing for the hearing-impaired: successes, failures, and implications for speech mechanisms / B. Moore // Speech Communication. — 2003. — Vol. 41, no. 1. — P. 81–91.
20. Comparison of adaptive noise reduction algorithms in dual microphone hearing aids / J.B. Maj, L. Royackers, J. Wouters, M. Moonen // Speech Communication. — 2006. — Vol. 48. — P. 957–979.
21. Griffiths, L. An alternative approach to linearly constrained adaptive beamforming / L. Griffiths, C. Jim // IEEE Transaction on Antennas Propag. — 1982. — Vol. AP-30, no. 1. — P. 27–34.
22. Frequency-domain criterion for the speech distortion weighted multichannel wiener filter for robust noise reduction / S. Doclo, A. Spriet, J. Wouters, M. Moonen // Speech Communication. — 2007. — Vol. 49, no. 7. — P. 636–656.

23. Kuo S. Active noise control: A tutorial review / S. Kuo, D. Morgan // Proc. IEEE. — 1999. — Vol. 87, no. 6. — P. 943–973.
24. Integrated active noise control and noise reduction in hearing aids / R. Serizel, M. Moonen, J. Wouters, S. Jensen // IEEE Trans. Audio, Speech, Lang. Process. — 2010. — Vol. 18, no. 6. — P. 1137–1146.
25. Kampre S.D. The influence of venting on the occlusion effect / S.D. Kampre, M.K. Wynne // The hearing journal. — 1996. — Vol. 49, no. 4. — P. 59–66.
26. Kates J.M. Feedback cancellation in hearing aids: results from a computer simulation / J.M. Kates // IEEE Transactions on Signal Processing. — 1991. — Vol. 39, no. 3. — P. 553–562.
27. Maxwell J.A. Reducing acoustic feedback in hearing aids / J.A. Maxwell, P.M. Zurek // IEEE Transactions on Speech and Audio Processing. — 1995. — Vol. 3, no. 4. — P. 304–313.
28. Chi H.-F. A novel approach of adaptive feedback cancellation for hearing aids / H.-F. Chi, S.X. Gao, S.D. Soli // Proceedings of IEEE International Symposium on Circuits and Systems. — USA, Orlando: 1999. — P. 195–198.
29. Hellgren, J. Analysis of feedback cancellation in hearing aids with filtered-x LMS and direct method of closed loop identification / J. Hellgren // IEEE Transactions on Speech and Audio Processing. — 2002. — Vol. 10, no. 2. — P. 119–131.
30. Combined feedback and noise suppression in hearing aids / A. Spriet, G. Rombouts, M. Moonen, J. Wouters // IEEE Transactions on Speech and Audio Processing. — 2007. — Vol. 15, no. 6. — P. 1777–1790.
31. Hellgren J. Variations in the feedback of hearing aids / J. Hellgren, F. Urban // Journal of the Acoustical Society of America. — 1999. — Vol. 106, no. 5. — P. 2821–2833.
32. Stinson M.R. Effect of handset proximity on hearing aid feedback / M.R. Stinson, G.A. Daigle // Journal of the Acoustical Society of America. — 2003. — Vol. 115, no. 3. — P. 1147–1156.
33. Kates J. Room reverberation effects in hearing aid feedback cancellation / J. Kates // Journal of the Acoustical Society of America. — 2000. — Vol. 109, no. 1. — P. 553–562.

34. Signal processing in high-end hearing aids: state of the art, challenges, and future trends / V. Hamacher, J. Chalupper, J. Eggers et al. // EURASIP Journal on Applied Signal Processing. — 2005. — no. 18. — P. 2915–2929.

35. Tan M.T. A novel low-voltage low-power wave digital filter bank for an intelligent noise reduction digital hearing instrument / M.T. Tan, J.S. Chang, Y.C. Tong // Proceedings of IEEE International Symposium on Circuits and Systems. — Australia, Sydney: 2001. — P. 681–684.

36. Вашкевич М.И. Неравнополосные банки фильтров для слуховых аппаратов: анализ алгоритмов, автоматизация проектирования / М.И. Вашкевич, А.А. Петровский // Автоматизация проектирования дискретных систем : материалы 7-й междунар. конф., Беларусь, Минск, 16-17 ноября 2010 г. — Минск, 2010. — С. 53–60.

37. Блейхут Р. Быстрые алгоритмы цифровой обработки сигналов / Блейхут Р. — М. : Мир, 1989. — 448 с.

38. Вашкевич М.И. Процессоры слуховых аппаратов для повышения разборчивости речи: проблемы, задачи исследования / М.И. Вашкевич, А.А. Петровский // Физика и радиоэлектроника в медицине и экологии: труды 9-й международной научно-технической конференции, Россия, Владимир, 29 июня — 2 июля 2010 г. — Владимир, 2010. — С. 235–238.

39. A digital master hearing aid / H. Levitt, A. Neuman, R. Mills, T. Schwander // Journal of Rehabilitation Research and Development. — 1986. — Vol. 23, no. 1. — P. 79–87.

40. Digital signal processing application for multiband loudness correction digital hearing aids and cochlear implants / N. Dillieer, T. Frölich, M. Kompis et al. // Journal of Rehabilitation Research and Development. — 1993. — Vol. 30, no. 1. — P. 95–109.

41. Chong K.S. A low energy FFT/IFFT processor for hearing aids / K.S. Chong, B.H. Gwee, J.S. Chang // Proceedings of IEEE International Symposium on Circuits and Systems. — USA, New Orleans: 2007. — P. 1169–1172.

42. Рабинер Л. Теория и применение цифровой обработки сигналов / Рабинер Л., Гоулд Б. — М. : Мир, 1978. — 848 с.

43. An ultra low-power programable DSP system for hearing aids and other audio application / T. Schneider, R. Brennan, P. Balsiger, A. Heubi // Proceedings of the International Conference on Signal Processing Applications and Technology. — Florida, USA: 1999. — P. 1–4.

44. A hearing aid simulator to test adaptive signal processing algorithms / R. Vicen-Bueno, R. Gil-Pita, M. Utrilla-Manso, L. Alvarez-Perez // Proceedings of the IEEE International Symposium on Intelligent Signal Processing. — Madrid, Spain: 2007. — P. 1–6.

45. Smith III J. Bark and ERB bilinear transforms / J. Smith III, J. Abel // IEEE Transactions on Speech and Audio Processing. — 1999. — Vol. 7, no. 6. — P. 697–708.

46. Vashkevich M. Low-delay hearing aid based on cochlear model with nonuniform subband acoustic feedback cancellation / M. Vashkevich, E. Azarov, A. Petrovsky // Proceedings of European Signal Processing Conference (EUSIPCO'2012). — Bucharest, Romania: 2012. — 28–31 August. — P. 514–518.

47. Вашкевич М.И. Неравнополосный косинусно-модулированный банк фильтров для аппроксимации шкалы барков / М.И. Вашкевич, А.А. Петровский // Доклады белорусского государственного университета информатики и радиоэлектроники. — 2009. — Т. 42, № 4. — С. 5–10.

48. Parfieniuk M. Tunable non-uniform filter bank mixing cosine modulation with perceptual frequency warping by allpass transformation / M. Parfieniuk, A.A. Petrovsky // Automatic Control and Computer Sciences. — 2004. — Vol. 38, no. 4. — P. 44–52.

49. Haykin S. Adaptive Filter Theory, 3rd Edition / S. Haykin. — Ney Jersey: Prentice Hall, 1996. — 936 p.

50. Petraglia M.R. Nonuniform subband adaptive filtering with critical sampling / M.R. Petraglia, P.B. Batalheiro // IEEE Transactions on Signal Processing. — 2008. — Vol. 56, no. 2. — P. 565–575.

51. Вашкевич М.И. Подавление эффекта акустической обратной связи в слуховых аппаратах с использованием неравнополосного банка фильтров / М.И. Вашкевич, И.С. Азаров, А.А. Петровский // Информатика. — 2012. — Т. 34.

52. Вайдьянатхан П.П. Цифровые фильтры, блоки фильтров и полифазные цепи с многочастотной дискретизацией : методический обзор / П.П. Вайдьянатхан // ТИИЭР: пер. с англ. — 1990. — no. 3. — P. 77–119.
53. Крот А.М. Теория анализа и синтеза бэнк-фильтров и их применение / А.М. Крот, В.О. Кудрявцев // Успехи современной радиоэлектроники. — 1999. — № 2. — С. 3–17.
54. Малла С. Вэйвлеты в обработке сигналов / Малла С. — М. : Мир, 2005. — 672 с.
55. Piotrowski A. Digital filter banks: analysis, synthesis and implementation for multimedia systems / A. Piotrowski, M. Parfieniuk. — Bialystok: Wydawnictwo Politechniki Bialostockiej, 2006. — 389 p.
56. Вашкевич М.И. Введение в цифровые банки фильтров / М.И. Вашкевич, А.А. Петровский// Анализаторы речевых и звуковых сигналов: методы, алгоритмы и практика (с МАТЛАВ примерами) под редакцией д-ра техн. наук профессора А.А. Петровского – Минск: Бестпринт, 2009. — С. 25–48.
57. Petrovsky Al. Real-time wavelet packet-based low bit rate audio coding on a dynamic reconfigurable system / Petrovsky Al., Krahe D., Petrovsky A.A. // 114th Audio Engineering Society Convention. — Netherlands, Amsterdam: 2003. — May. — P. 1–22.
58. Daubechies I. Factoring wavelet transforms into lifting steps / Daubechies I., Sweldens W. // The Journal of Fourier Analysis and Applications. — 1998. — Vol. 4, no. 2. — P. 247–269.
59. Princen J. The design of nonuniform modulated filterbanks / Princen J. // IEEE Transaction on Signal Processing. — 1995. — Vol. 43, no. 11. — P. 2550–2560.
60. Li J. A simple design method for near-perfect-reconstruction nonuniform filter banks / J. Li, T. Q. Nguyen, S. Tantaratana // IEEE Transactions on Signal Processing. — 1997. — Vol. 45, no. 8. — P. 2105–2109.
61. Galijasevic E. Design of allpass-based non-uniform oversampled DFT filter banks / E. Galijasevic, J. Kliewer // Proceedings of International Conference on Acoustics, Speech, and Signal Processing. — 2002. — P. 1181–1184.

62. Gülzow T. Comparison of a discrete wavelet transformation and a nonuniform polyphase filterbank applied to spectral-subtraction speech enhancement / T. Gülzow, A. Engelsberg, U. Heute // Signal Processing. — 1998. — Vol. 64, no. 1. — P. 5–19.

63. Koilpillai R. Cosine-modulated FIR filter banks satisfying perfect reconstruction / R. Koilpillai, P. Vaidyanathan // IEEE Transactions on Signal Processing. — 1992. — Vol. 40, no. 4. — P. 770–783.

64. Rothweiler J. Polyphase quadrature filters — a new subband technique / J. Rothweiler // Proceedings of International Conference on Acoustics, Speech, and Signal Processing. — USA, Boston: 1983. — P. 1280–1283.

65. Koilpillai R. A spectral factorization approach to pseudo-QMF design / R. Koilpillai, P. Vaidyanathan // IEEE Transactions on Signal Processing. — 1993. — Vol. 41, no. 1. — P. 82–92.

66. Goh C.-K. An efficient algorithm for the design of weighted minimax M-channel cosine-modulated filter banks / C.-K. Goh, Y.-C. Lim // IEEE Transactions on Signal Processing. — 1998. — Vol. 46, no. 5. — Pp. 1426–1430. — P. 1426–1430.

67. Lin Y. Kaiser window approach for the design of prototype filters of cosine-modulated filterbanks / Y. Lin, P. Vaidyanathan // IEEE Signal Processing Letters. — 1998. — Vol. 5, no. 6. — P. 132–134.

68. Вашкевич М.И. Обобщенная полифазная структура косинусно-модулированного банка фильтров / М.И. Вашкевич, А.А. Петровский // Цифровая обработка сигналов и ее применение: труды 13-й междунар. конф., Россия, Москва, 30 марта – 1 апреля, 2011 г. — Москва, 2011. — С. 91–94.

69. Оппенгейм А. Вычисление спектров с неравномерным разрешением посредством быстрого преобразования Фурье / А. Оппенгейм, Д. Джонсон, К. Штейглиц // ТИИЭР: пер. с англ. — 1971. — Т. 59, № 2. — С. 138–139.

70. Вашкевич М.И. К вопросу о выборе коэффициентов децимации в неравнополосных банках фильтров / М.И. Вашкевич, А.А. Петровский // Современные средства связи: тезисы докладов международной

научно-технической конференции, Беларусь, Минск, 7–9 октября 2008 г. – Минск, 2008. — С. 139.

71. Vaughan R. The theory of bandpass sampling / R. Vaughan, N. Scott, R. White // IEEE Transactions on signal processing. — 1991. — Vol. 39, no. 9. — P. 1973–1984.

72. Parfieniuk M. Simple rule of selection of subsampling ratios for warped filter banks / M. Parfieniuk, A. Petrovsky // Известия Белорусской инженерной академии. — 2003. — № 1. — С. 130–134.

73. Parfieniuk M. Reduced complexity synthesis part of non-uniform near-perfect-reconstruction DFT filter bank based on all-pass transformation / M. Parfieniuk, A. Petrovsky // Proceedings of European conference on Circuit Theory and Design. — Poland, Cracow: 2003. — P. 5–8.

74. Парфенюк М. Неравнополосный банк фильтров с фазовым преобразованием и объединением субполос для обработки речевых / М. Парфенюк, М.И. Вашкевич, А.А. Петровский // Речевые технологии. — 2009. — № 4. — С. 53–69.

75. Вашкевич М.И. Проектирование передискретизированного неравнополосного косинусно-модулированного банка фильтров / М.И. Вашкевич, А.А. Петровский // Информатика. — 2011. — № 2(30). — С. 21–39.

76. Vashkevich M. A design method for oversampled warped cosine-modulated filter banks / M. Vashkevich, A. Petrovsky // Proceedings of the International Conference Signals and electronic systems, Gliwice, Poland, 7–10 Sept., 2010. — Gliwice, 2010. — P. 65–68.

77. Vashkevich M.I. Practical design of multi-channel oversampled warped cosine-modulated filter banks / M.I. Vashkevich, W. Wan, A.A. Petrovsky // Proceedings of the International Communication Conference on Wireless Mobile and Computing, Shanghai, China, 14–16 November, 2011. — Shanghai, 2011. — P. 44–49.

78. A novel approach to the design of oversampling complex-modulated digital filter banks / C. Stöcker, T. Kurbiel, D. Alfsmann, H.G. Göckler // Proceedings of European Signal Processing Conference (EUSIPCO'2009). — Glasgow, Scotland: 2009. — P. 2648–2652.

79. Crochiere R.E. Multirate Digital Signal Processing / R.E. Crochiere, L.R. Rabiner. — Englewood Cliffs, NJ, USA: Prentice Hall, 1983. — 441 p.

80. Kuo S.M. Integrated frequency-domain digital hearing aid with the lapped transform / Kuo S.M., Voepel S. // Electronic letters. — 1992. — Vol. 28, no. 23. — P. 2117–2118.

81. Malvar H.S. The LOT: transform coding without blocking effects / H.S. Malvar, D.H. Staelin // IEEE Trans. on Acoustics, Speech, Signal Processing. — 1989. — Vol. ASSP-37, no. 4. — P. 553–559.

82. Jayant N.S.. Digital coding of waveforms: principles and applications to speech and video / N.S. Jayant, P. Noll. — Prentice Hall, 1984. — 688 p.

83. Wyrsh S. Subband signal processing for hearing aids / S. Wyrsh, Kaelin A. // Proceedings of the IEEE International Symposium on Circuits and System, Orlando, USA, 30 May – 2 Jun, 1999. — Orlando, 1999. — Vol. 3. — P. 29–32.

84. Wyrsh S. A dsp implementation of a digital hearing aid with recruitment of loudness compensation and acoustic echo cancellation / S. Wyrsh, A. Kaelin // IEEE Workshop on Applications of Signal Processing in Audio and Acoustics, New Paltz, NY, USA, 19–22 Oct., 1997. — New Paltz, 1997.

85. Kaelin A. A digital frequency-domain implementation of a very high gain hearing aid with compensation for recruitment of loudness and acoustic echo cancellation / A. Kaelin., A. Lindgren, S. Wyrsh // Signal Processing. — 1998. — no. 64. — P. 71–85.

86. Wyrsh S. Adaptive feedback cancelling in subbands for hearing aids / S. Wyrsh, A. Kaelin // Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Phoenix, USA, 15–19 March, 1999. — Phoenix, 1999. — Vol. 2. — P. 921–924.

87. Вашкевич М.И. Исследование возможности применения модифицированного дискретного косинусного преобразования с деформацией частотной оси для построения слухового аппарата / М.И. Вашкевич, М. Парфенюк, А.А. Петровский // Цифровая обработка сигналов и ее применение

ние: труды 15-й междунар. конф., Россия, Москва, 27–29, 2013 г. – Москва, 2013. — С. 109–112.

88. Püschel M. Algebraic signal processing theory: Cooley-Tukey type algorithms for DCTs and DSTs / M. Püschel, J.M.F. Moura // IEEE Transaction on Signal Processing. — 2008. — Vol. 56, no. 4. — P. 1502–1521.

89. Вашкевич М.И. Применение полиномиальных алгебр и теории Галуа для синтеза быстрых алгоритмов дискретных косинусных преобразований / М.И. Вашкевич, А.А. Петровский // Цифровая обработка сигналов. — 2011. — № 3. — С. 2–10.

90. Вашкевич М.И. Алгебраический метод синтеза быстрых алгоритмов дискретного косинусного преобразования произвольного размера / М.И. Вашкевич, А.А. Петровский // Автоматика и вычислительная техника. — 2012. — № 5. — С. 48–57.

91. Vashkevich M. A low multiplicative complexity fast recursive DCT-2 algorithm / M. Vashkevich, A. Petrovsky // Computing Research Repository [Electronic resource]. — 2012. — Mode of access: <http://arxiv.org/pdf/1203.3442v1.pdf>. — Date of access: 26.07.2012.

92. Миано Дж. Форматы и алгоритмы сжатия изображений в действии / Дж. Миано — М. : Триумф, 2003. — 336 с.

93. Britanak V. Discrete Cosine and Sine Transforms: General Properties, Fast Algorithms and Integer Approximations / V. Britanak, P.C. Yip, K.R. Rao. — UK, Oxford: Academic, 2007. — 368 p.

94. Чернов В.М. Арифметические методы синтеза быстрых алгоритмов дискретных ортогональных преобразований / В.М. Чернов— М.: Физматлит, 2007. — 264 с.

95. Chen W.-H. A fast computational algorithm for the discrete cosine transform / W.-H. Chen, C. Smith, S. Fralick // IEEE Transactions on Communications. — 1977. — Vol. 25, no. 9. — P. 1004–1009.

96. Arai Y. A fast DCT-SQ scheme for images / Y. Arai, T. Agui, M. Nakajima // IEICE Transactions. — 1988. — Vol. 71, no. 11. — P. 1095–1097.

97. Loeffler C. Practical fast 1-D DCT algorithms with 11 multiplications / C. Loeffler, A. Ligtenberg, G.S. Moschytz // Proceedings

of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Glasgow, UK, 23–26 May, 1989. — Glasgow, 1989. — P. 988–991.

98. Liang J. Fast multiplierless approximations of the DCT with the lifting scheme / J. Liang, T.D. Tran // IEEE Transaction on Signal Processing. — 2001. — Vol. 49. — P. 3032–3044.

99. Parfieniuk M. Structurally orthogonal finite precision implementation of the eight point DCT / M. Parfieniuk, A. Petrovsky // Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Toulouse, France, 23–26 May, 1989. — Toulouse, 2006. — Vol. 3. — 2006. — P. 936–939.

100. Vashkevich M. FPGA implementation of short critical path CORDIC-based approximation of the eight-point DCT / M. Vashkevich, M. Parfieniuk, A. Petrovsky // Proceedings of the International Conference Pattern Recognition and Information Processing, Minsk, Belarus, 19–21 May, 2009. — Minsk, 2009. — P. 161–164.

101. SPIRAL: Code generation for DSP transforms / M. Püschel, J.M.F. Moura, J. Johnson et al. // Proceedings of the IEEE, special issue on “Program Generation, Optimization, and Adaptation”. — 2005. — Vol. 93, no. 2. — P. 232–275.

102. A systematic approach for synthesizing vlsi architectures of lifting-based filter banks and transforms / R. Bartholomä, T. Greiner, F. Kesel, W. Rosenstiel // IEEE Transaction on Circuits and Systems. — 2008. — Vol. 55-I, no. 7. — P. 1939–1952.

103. Белый А.А. Алгоритмы быстрого преобразования Фурье и их свойства / А.А. Белый, Е.И. Бовбель, В.И. Микулович // Зарубежная радиоэлектроника. — 1979. — № 2. — С. 3–29.

104. Вайрадян А.С. Алгоритмы вычисления цифровых сверток / А.С. Вайрадян, И.П. Пчелинцев, М.М. Челышев // Зарубежная радиоэлектроника. — 1982. — № 3. — С. 3–34.

105. Трахтман А.М. Основы теории дискретных сигналов на конечных интервалах / А.М. Трахтман, В.А. Трахтман— М.: Советское радио, 1975. — 208 с.

106. Нуссбаумер Г. Быстрое преобразование Фурье и алгоритмы вычисления свертков / Г. Нуссбаумер — М.: Радио и связь, 1985. — 248 с.

107. Крот А.М. Дискретные модели динамических систем на основе полиномиальной алгебры / А.М. Крот — Мн.: Наука и техника, 1985. — 312 с.

108. Püschel, M. The algebraic approach to the discrete cosine and sine transforms and their fast algorithms / M. Püschel, J.M.F. Moura // SIAM Journal on Computing. — 2003. — Vol. 32, no. 5. — P. 1280–1316.

109. Лабунец В.Г. Алгебраическая теория сигналов и систем: цифровая обработка сигналов / Лабунец В.Г. — Красноярск: Издательство Красноярского университета, 1984. — 243 с.

110. Воеводин В.В. Линейная алгебра / В.В. Воеводин — М.: Наука, 1980. — 400 с.

111. Ахо А. Построение и анализ вычислительных алгоритмов / А. Ахо, Дж. Хопкрофт, Дж. Ульман — М.: Мир, 1979. — 536 с.

112. Ван дер Варден Б.Л. Алгебра / Б.Л. Ван дер Варден — М.: Наука, 1978. — 623 с.

113. Wang Z. Fast algorithm for the discrete W-transform and for the discrete Fourier transform / Z. Wang // IEEE Transactions on Acoustics, Speech, and Signal Processing. — 1984. — Vol. 32, no. 4. — P. 803–816.

114. Lee B. A new algorithm to compute the discrete cosine transform / B. Lee // IEEE Transactions on Acoustics, Speech, and Signal Processing. — 1984. — Vol. 32, no. 6. — P. 1243–1245.

115. Hou H. A fast recursive algorithm for computing the discrete cosine transform / H. Hou // IEEE Transactions on Acoustics, Speech, and Signal Processing. — 1987. — Vol. 35, no. 10. — P. 1455–1461.

116. Оппенгейм, А. Цифровая обработка сигналов / А. Оппенгейм, Р. Шафер — М.: Техносфера, 2006. — 856 с.

117. Стренг Г. Линейная алгебра и её применения / Г. Стренг — М.: Мир, 1980. — 447 с.

118. Gustafsson S. A novel psychoacoustically motivated audio enhancement algorithm preserving background noise characteristics / S. Gustafsson, P. Jax, P. Vary // Proceedings of the IEEE International 200

Conference on Acoustics, Speech, and Signal Processing, Washington, USA, 12–15 May, 1998. — Washington, 1998. — P. 397-400.

119. Petrovsky A. Warped DFT based perceptual noise reduction system / A. Petrovsky, M. Parfieniuk, A. Borowicz // Proceedings of the 116th Audio Engineering Society Convention, Berlin, Germany, 8–11 May, 2004. — Berlin, 2004. — P. 1-16.

120. Дженкинс Г. Спектральный анализ и его приложения / Г. Дженкинс, Д. Ваттс — М.: Мир, 1971. — 283 с.

121. Cohen I. Noise estimation by minima controlled recursive averaging for robust speech enhancement / I. Cohen // IEEE Signal Processing Letters. — 2002. — Vol. 9, no. 1. — P. 12–15.

122. Петровский А.А. Дискретное преобразование Фурье с неравномерным частотным разрешением в перцептуальных системах редактирования шума в речи / А.А Петровский, А. Борович, М. Парфенюк // Речевые технологии. — 2008. — № 3. — С. 16–26.

123. Kamm C.A. Speech recognition and the articulation index for normal and hearing-impaired listeners / C.A. Kamm, D.D. Dirks, T.S. Bell // Journal of the Acoustical Society of America. — 1985. — Vol. 77, no. 1. — P. 281–288.

124. An integrated approach for noise reduction and dynamic range compression in hearing aids / K. Ngo, S. Doclo, A. Spriet et al. // Proceedings of the European Signal Processing Conference, Lausanne, Switzerland, 25–29 August, 2008. — Lausanne, 2008.

125. Вашкевич М.И. Компрессор речевых сигналов для слуховых аппаратов на основе кохлеарного банка фильтров / М.И. Вашкевич, А.А. Петровский // Цифровая обработка сигналов и ее применение: труды 14-й междунар. конф., Россия, Москва, 28–30, 2012 г. — Москва, 2012. — С. 87–91.

126. Acoustic echo control. An application of very-high-order adaptive filters / C. Breining, P. Dreiscitel, E. Hansler et al. // IEEE Signal Processing Magazine. — 1999. — Vol. 16, no. 4. — P. 42–69.

127. Acoustic feedback reduction based on filtered-X LMS and normalized filtered-X LMS algorithms in digital hearing aids based on

WOLA filterbank / R. Vicen-Bueno, A. Martinez-Leira, R. Gil-Pita, M. Rosa-Zurera // Proceedings of the IEEE International Symposium on Intelligent Signal Processing, Madrid, Spain, 3–5 October, 2007. — Madrid, 2007. — P. 1–6.

128. Hellgren, J. System identification of feedback in hearing aids / J. Hellgren, T. Lunner, S. Arlinger // Journal of the Acoustical Society of America. — 1991. — Vol. 105, no. 6. — P. 3481–3496.

129. ITU-T Rec. P 862. Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow band telephone networks and speech coders [Electronic resource]. — 2001. — Mode of access : <http://www.itu.int/rec/T-REC-P.862-200102-I/en>. — Date of access: 11.05.2012.

130. Farhang-Boroujeny B. Adaptive filters: theory and applications / B. Farhang-Boroujeny. — New York, USA: Wiley, 1998. — 548 p.

131. Agnew J. Just noticeable and objectionable group delays in digital hearing aids / J. Agnew, J.M. Thornton // Journal of the American Academy of Audiology. — 2000. — Vol. 11, no. 6. — P. 330–336.

132. Stone M.A. Tolerable hearing aids delays. ii: Estimation of limits imposed during speech production / M.A. Stone, B.C.J. Moore // Journal of the American Academy of Audiology. — 2002. — Vol. 11, no. 6. — P. 325–338.

133. Ryan J. A digital signal processor for musicians and audiophiles / J. Ryan, S. Tewari // Hearing Review. — 2009. — Vol. 16, no. 2. — P. 38–41.

134. Harteneck M. Design of near perfect reconstruction oversampled filter banks for subband adaptive filters / M. Harteneck, S. Weiss, R.W. Stewart // IEEE Transaction on Circuits and Systems. — 1999. — Vol. 46, no. 8. — P. 1081–1086.

135. Biological basis of hearing-aid design / M.B. Sachs, I.C. Bruce, R.L. Miller, E.D. Young // Annals of biomedical engineering. — 2002. — Vol. 30, no. 2. — P. 157–168.

136. Berger K. Prescription of hearing aids: rationale, procedure and results (5-th ed.) / K. Berger, E. Hagberg, R. Rane. — Kent, OH: Herald Publishing, 1988. — 63 p.

137. McCandless G.A. Prescription of gain/output (POGO) for hearing aids / G.A. McCandless, P.E. Lyregaard // Hearing Instruments. — 1983. — no. 34. — P. 16–21.

138. Byrne D. The national acoustic laboratories (NAL) new procedure for selecting the gain and frequency response of a hearing aid / D. Byrne, H. Dillon // Ear and Hearing. — 1986. — no. 7. — P. 257–265.

139. Byrne D. Modified hearing aid selection process for severe/profound hearing losses / D. Byrne, A. Parkinson, P. Newall // The Vanderbilt hearing aid report II. — 1991. — no. 7. — P. 293–300.

140. Oppenheim A. Digital signal processing / A. Oppenheim, R. Schaffer. — New Jersey: Prentice Hall, Englewood Cliffs, 1983. — 784 p.

141. Schaub A. Digital hearing aids / A. Schaub. — New-York: Thieme, 2008. — 190 p.

142. Simulation of hearing loss using compressive gammachirp auditory filters / H. Hu, J. Sang, M.E. Lutman, S. Bleek // Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Prague, Czech Republic, 22–27 May, 2011. — Prague, 2011. — P. 5428–5431.

143. Peled A. A new approach to the realization of nonrecursive digital filters / A. Peled, B. Lui // IEEE Transaction on Audio and Electroacoustics. — 1973. — Vol. AU-21, no. 6. — P. 477–485.

144. White S.A. Applications of distributed arithmetic to digital signal processing: A tutorial review / S.A. White // IEEE ASSP Magazine. — 1989. — Vol. 6. — P. 4–19.

145. Вашкевич М.И. Применение распределенной арифметики при аппаратной реализации неравнополосных банков фильтров / М.И. Вашкевич, А.А. Петровский // Информационные системы и технологии : материалы седьмой международной конференции, Беларусь, Минск, 16-17 ноября 2009 г. — Минск, 2009. — С. 198–201.

146. Vashkevich M. High-accuracy implementation of fast dct algorithm based on algebraic integer encoding / M. Vashkevich, M. Parfieniuk, A. Petrovsky // Proceedings of the International Conference Signals and electronic systems, Wroclaw, Poland, 18–21 Sept., 2012. — Wroclaw, 2012. — P. 65–68.

147. Parfieniuk M. Shortening the critical path in CORDIC-based approximations of the eight-point DCT / M. Parfieniuk // Proceedings of the International Conference on Signals and Electronic Systems, Krakow, Poland, 14–17 May, 2008. — Krakow, 2008. — P. 405–408.

148. Dimitrov V. Multiplierless DCT algorithm for image compression applications / V. Dimitrov, K. Wahid // International Journal: Information Theories and Applications. — 2004. — Vol. 11, no. 2. — P. 162–169.

149. Dimitrov V. On the error-free computation of fast cosine transform / V. Dimitrov, K. Wahid // International Journal: Information Theories and Applications. — 2005. — Vol. 12, no. 4. — P. 321–327.

150. Кнут Д.Э. Искусство программирования. Том 2. Получисленные алгоритмы / Д.Э. Кнут — М.: Вильямс, 2000. — 500 с.

151. Босс В. Теория групп / В. Босс — М.: Либроком, 2009. — 216 с.

152. Гросман И. Группы и их графы / И. Гросман, В. Магнус — М.: Мир, 1971. — 248 с.

153. Данилов, Ю.А. Многочлены Чебышева / Ю.А. Данилов — Мн.: Вышэйшая школа, 2009. — 160 с.

154. Wan W.G. A new solution for cochlear macromechanics / W.G. Wan, C.X. Fan // Acoustica. — 1991. — Vol. 75. — P. 79–82.

155. Wan W. A two-dimensional nonlinear cochlear model for speech processing: response to pure tone / W. Wan, A. Petrovsky, C. Fan // Proceedings of the 6th International Fase – Congress. — Zurich, Switzerland: 1992. — July. — P. 233–236.

156. Petrovsky A.A. A digital cochlear model as a base of anthropomorphic speech processing / A.A. Petrovsky, D.S. Likhachov // Proceedings of the International Conference on Neural Networks and Artificial Intelligence, Minsk, Belarus, 12–14 November, 2003. — Minsk, 2003. — P. 126–131.

157. Petrovsky A. The synthesis of high order digital bandpass filters with tunable center frequency and bandwidth / A. Petrovsky // Proceedings of the European Signal Processing Conference, Trieste, Italy, 10–13 September, 1996. — Trieste, 1996. — P. 1527–1530.

ПРИЛОЖЕНИЕ А

Краткие сведения из теории групп и теории Галуа

Определение группы. Группой G называется совокупность элементов, на которой задана *групповая (бинарная) операция* « \cdot », сопоставляющая любой паре элементов $g_1, g_2 \in G$ некоторый элемент g_3 из той же совокупности G . При этом групповая операция должна удовлетворять трем условиям [151]:

- ассоциативность: $g_1 \cdot (g_2 \cdot g_3) = (g_1 \cdot g_2) \cdot g_3$;
- существование единицы: в группе есть такой элемент id , что $\text{id} \cdot g = g \cdot \text{id} = g$ для всех $g \in G$;
- существование обратного элемента: для любого $g \in G$ существует такой $g^{-1} \in G$, что $g \cdot g^{-1} = g^{-1} \cdot g = \text{id}$.

Внутреннюю структуру некоторых групп можно описать с помощью их подгрупп. Слово «подгруппа» означает «группа внутри группы»; точнее, множество H называется подгруппой группы G , если [152]:

- каждый элемент множества H является элементом группы G ;
- H есть группа (относительно бинарной операции, определенной в группе G).

Поле разложение полинома. Пусть $p(x) \in \mathbb{F}[x]$, *полем разложения* \mathbb{E} полинома $p(x)$ считается наименьшее расширение поля \mathbb{F} , содержащее все корни $p(x)$. Например, $\mathbb{Q}_{\sqrt{2}}$ — поле разложения полинома $p(x) = x^2 - 2$. $\mathbb{Q}_{\sqrt{2}}$ образуется присоединением числа $\sqrt{2}$ к полю \mathbb{Q} .

Автоморфизмы полей. Взаимнооднозначное отображение поля \mathbb{E} на себя называется *автоморфизмом*. \mathbb{F} -автоморфизмом поля \mathbb{E} называется такой автоморфизм φ , для которого $\varphi(x) = x$, $\forall x \in \mathbb{F}$. Группу \mathbb{F} -автоморфизмов называют группой Галуа полинома $p(x)$ и обозначают $\text{Gal}(p)$ или $\text{Gal}(\mathbb{E}/\mathbb{F})$. В качестве примера определим функцию $\varphi: \mathbb{Q}_{\sqrt{2}} \rightarrow \mathbb{Q}_{\sqrt{2}}$ такую, что

$$\varphi(a + b\sqrt{2}) = a - b\sqrt{2}, \quad a, b \in \mathbb{Q},$$

тогда φ представляет собой \mathbb{Q} -автоморфизм поля $\mathbb{Q}_{\sqrt{2}}$. Группа автоморфизмов $\text{Gal}(p) = \{\text{id}, \varphi\}$ представляет собой циклическую группу второго порядка, где id — \mathbb{Q} -автоморфизм, который оставляет все элементы поля $\mathbb{Q}_{\sqrt{2}}$ на своих местах. Очевидно, что $\varphi \cdot \varphi = \text{id}$.

Важным результатом теории Галуа является установление связи между структурой подполей поля разложения полинома $p(x)$ и структурой подгрупп группы Галуа $\text{Gal}(\mathbb{E}/\mathbb{F})$ [151]. Каждой подгруппе H группы $\text{Gal}(\mathbb{E}/\mathbb{F})$ отвечает подполе $\mathbb{L} \subset \mathbb{E}$, состоящее из элементов \mathbb{E} , неподвижных под действием автоморфизмов из H . Аналогично, каждому подполю $\mathbb{L} \subset \mathbb{E}$ отвечает подгруппа H группы Галуа, оставляющая элементы из \mathbb{L} на месте. В результате изучение всех подполей поля \mathbb{E} сводится к изучению всех подгрупп группы $\text{Gal}(\mathbb{E}/\mathbb{F})$. При этом каждой башне (цепочке вложенных) полей

$$\mathbb{F} = \mathbb{L}_0 \subset \mathbb{L}_1 \subset \cdots \subset \mathbb{L}_r = \mathbb{E} \quad (\text{A.1})$$

отвечает нормальный ряд вложенных (в обратном направлении) групп

$$\text{Gal}(\mathbb{E}/\mathbb{F}) = G_0 \supset G_1 \supset \cdots \supset G_r = \{1\}$$

и наоборот (соответствие Галуа).

Таким образом, соответствие Галуа позволяет определить все подполя поля разложения полинома $p(x)$. В каждом подполе $p(x)$ раскладывается единственным образом в произведение неприводимых над этим подполем полиномов. Используя (A.1), можно выполнить поэтапную факторизацию $p(x)$, которая необходима для синтеза быстрого алгоритма ДКП.

ПРИЛОЖЕНИЕ Б

Полиномы Чебышева

Полиномы Чебышева [153] образуют специальный класс ортогональных многочленов и играют важную роль во многих областях математики. В данном приложении рассматриваются только те свойства полиномов Чебышева, которые используются при синтезе быстрых алгоритмов ДКП.

Обозначим через $C_0(x) = 1$ и $C_1(x)$ полиномы нулевой и первой степени, тогда $C_n(x)$ для $n > 1$ определяется рекуррентной формулой

$$C_n(x) = 2xC_{n-1}(x) - C_{n-2}(x).$$

Данный ряд полностью определяется условием $C_0 = 1$ и выбором C_1 . Для синтеза быстрых алгоритмов ДКП и ДСП практическую значимость имеют четыре частных случая полиномов Чебышева [108]. Они обозначаются как $C \in \{T, U, V, W\}$ и называются полиномами Чебышева первого, второго, третьего и четвертого рода (таблица Б.1).

Таблица Б.1. Полиномы Чебышева 1–4 рода

	Первые члены ряда	Аналитический вид ($\cos \theta = x$)	Тип симметрии	Корни ($0 \leq k < n$)
T_n	$1, x$	$\cos(n\theta)$	$T_{-n} = T_n$	$\cos \frac{(k+\frac{1}{2})\pi}{n}$
U_n	$1, 2x$	$\frac{\sin(n+1)\theta}{\sin \theta}$	$U_{-n} = -U_{n-2}$	$\cos \frac{(k+1)\pi}{n+1}$
V_n	$1, 2x - 1$	$\frac{\cos(n+\frac{1}{2})\theta}{\cos \frac{1}{2}\theta}$	$V_{-n} = V_{n-1}$	$\cos \frac{(k+\frac{1}{2})\pi}{n+\frac{1}{2}}$
W_n	$1, 2x + 1$	$\frac{\sin(n+\frac{1}{2})\theta}{\sin \frac{1}{2}\theta}$	$W_{-n} = -W_{n-1}$	$\cos \frac{(k+1)\pi}{n+\frac{1}{2}}$

ПРИЛОЖЕНИЕ В

Кохлеарный банк фильтров

Для выполнения частотного анализа и разбиения сигнала на субполосные составляющие в слуховом аппарате может быть использован банк кохлеарных фильтров, основанный на SDCM-модели (*Second order Difference Cochlear Model* — разностная кохлеарная модель второго порядка) [46, 125]. Согласно модели функционирование улитки уха на электрическом уровне описывается разностным уравнением [154–156]

$$y_k[n] + b_{1,k}y_k[n-1] + b_{2,k}y_k[n-2] = A_k a_{0,k} (u_s[n] - u_s[n-2]), \quad (\text{B.1})$$

где u_s — входной синусоидальный сигнал, характеризующий скорость перемещения стремечка в ухе человека;

k — номер сегмента базилярной мембраны после дискретизации;

$y_k[n]$ — перемещение или так называемая пучность базилярной мембраны в позиции x_k ;

$b_{1,k}$, $b_{2,k}$, $a_{0,k}$ и A_k — параметры, определяемые исходя из физических характеристик базилярной мембраны, таких как жесткость, затухание, длина, масса и др. в позиции x_k .

Модель (B.1) можно представить в виде банка полосовых фильтров с высокой степенью перекрытия полос пропускания. Получаемый банк обладает свойством неравнополосности и согласован с работой улитки уха человека. Передаточная функция k -го кохлеарного фильтра имеет вид

$$H_k(z) = A_k \frac{a_{0k}(1 - z^2)}{1 + b_{1k}z^{-1} + b_{2k}z^{-2}}.$$

Центральная частота фильтра (ω_{0k}) и полоса пропускания по уровню 3 дБ ($\Delta\omega_{0k}$) выражаются следующим образом:

$$\cos \omega_{0k} = -b_{1k}/(1 + b_{2k}); \quad \Delta\omega_{0k} = 2(1 - b_{2k})/(1 + b_{2k}).$$

Используя подход, предложенный в [157], кохлеарный банк фильтров может быть эффективно реализован в виде набора полосовых фильтров с

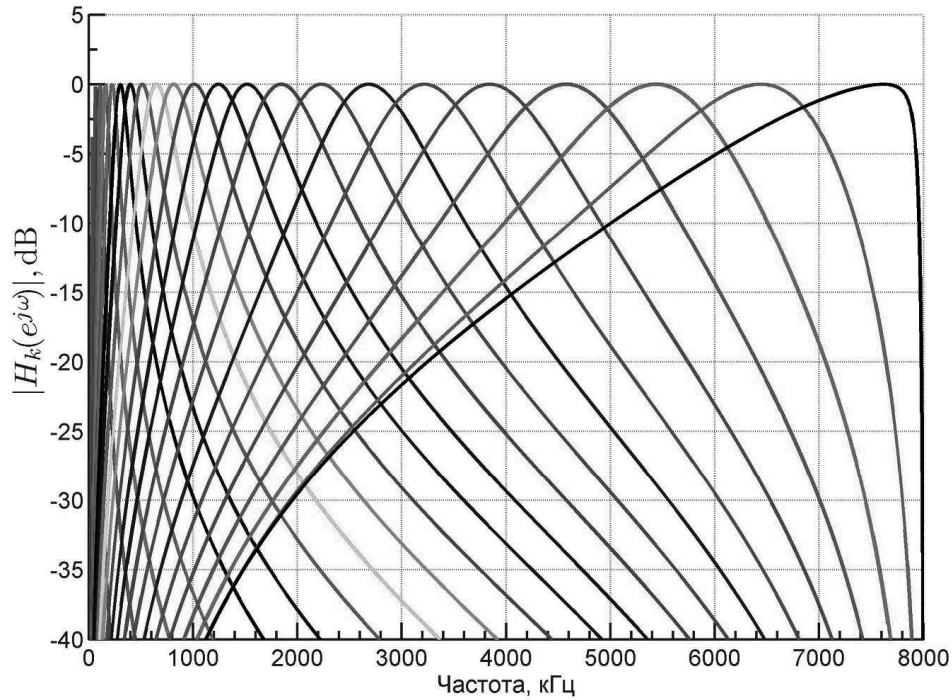


Рис. В.1. АЧХ кохлеарного банка фильтров

перестраиваемой центральной частотой и полосой пропускания. Передаточная функция такого фильтра определяется как

$$H(z) = a_0 \frac{1 - z^{-2}}{1 + (a_0 - 1)gz^{-1} + (1 - 2a_0)z^{-2}}.$$

Коэффициенты a_0 и g зависят только от ширины полосы пропускания фильтра $\Delta\omega$ и центральной частотой ω_0 соответственно:

$$a_0 = \frac{\Delta\omega T}{2 + \Delta\omega T}, \quad g = 2 \cos \omega_0 T,$$

где T – период дискретизации.

АЧХ 22-канального кохлеарного банка фильтров показана на рис. В.1. Групповая задержка в банке фильтров колеблется в пределах 0,8–4 мс.

Кохлеарный банк фильтров реализуется в виде набора полосовых фильтров с перестраиваемой центральной частотой и полосой пропускания. Передаточная функция такого фильтра определяется как

$$H(z) = a_0 \frac{1 - z^{-2}}{1 + (a_0 - 1)gz^{-1} + (1 - 2a_0)z^{-2}}, \quad (\text{В.2})$$

где коэффициенты a_0 и g определяются полосой пропускания и центральной частотой, соответственно.

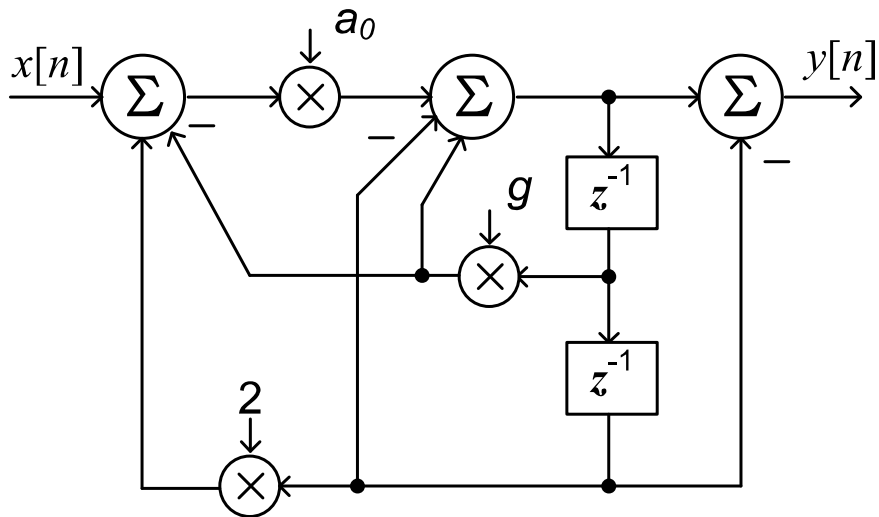


Рис. В.2. Схема перестраиваемого полосового фильтра

В [157] на основании передаточной функции (В.2) записывалась следующая система уравнений, описывающая работу полосового фильтра:

$$\begin{cases} w[n] = x[n] - gv[n - 1] + 2v[n - 2]; \\ v[n] = a_0w[n] + gv[n - 1] - v[n - 2]; \\ y[n] = v[n] - v[n - 2]. \end{cases} \quad (\text{В.3})$$

Соответствующая схема фильтра показана на рис. В.2. В структуре фильтра имеется два элемента памяти для хранения отсчетов $v[n - 1]$ и $v[n - 2]$. Вычисление каждого выходного отсчета $y[n]$ требует двух операций умножения и пяти сложений.