

Афанасьев А. А.
Рыболовлев А. А.
Рыжков А. П.

ЦИФРОВАЯ ОБРАБОТКА СИГНАЛОВ

*Рекомендовано Федеральным государственным бюджетным
образовательным учреждением высшего профессионального образования
«Московский технический университет связи и информатики»
в качестве учебного пособия для студентов высших учебных заведений,
обучающихся по укрупненной группе направлений подготовки
11.00.00 – «Электроника, радиотехника и системы связи»
Регистрационный номер рецензии 3107 от 01.09.2015 г. МГУП*

Москва
Горячая линия – Телеком
2019

УДК 621.391.037.37

ББК 32.811.3

A94

Рецензенты: доктор техн. наук, профессор *А. С. Аджемов*;
канд. техн. наук, профессор *В. С. Сперанский*

Афанасьев А. А., Рыболовлев А. А., Рыжков А. П.

A94 Цифровая обработка сигналов. Учебное пособие для вузов. –
М.: Горячая линия – Телеком, 2019. – 356 с.: ил.

ISBN 978-5-9912-0611-2.

Изложены базовые теоретические и практические вопросы в области анализа, синтеза и экспериментального исследования систем цифровой обработки сигналов (ЦОС). Приведены основные методы расчета характеристик устройств ЦОС, показаны особенности их анализа в различных условиях функционирования, даны основные подходы к синтезу систем ЦОС, рассмотрены основные происходящие в них физические процессы. Изложенные теоретические сведения сопровождаются практическими примерами решения прикладных задач.

Для студентов вузов, обучающихся по укрупненной группе направлений подготовки 11.00.00 – «Электроника, радиотехника и системы связи», будет полезно для специалистов, занимающихся проектированием систем ЦОС.

ББК 32.811.3

Адрес издательства в Интернет WWW.TECHBOOK.RU

Учебное издание

Афанасьев Андрей Алексеевич, Рыболовлев Александр Аркадьевич,
Рыжков Александр Павлович

ЦИФРОВАЯ ОБРАБОТКА СИГНАЛОВ

Учебное пособие для вузов

Редактор Ю. Н. Чернышов
Компьютерная верстка Ю. Н. Чернышова
Обложка художника О. Г. Карповой

Подписано в печать 12.11.2016. Формат 60×90/16. Усл. печ. л. 22,25.

Тираж 1000 экз. (3 завод. – 50 экз.). Изд. № 160611

ООО «Научно-техническое издательство «Горячая линия – Телеком»

ISBN 978-5-9912-0611-2

© А. А. Афанасьев, А. А. Рыболовлев,
А. П. Рыжков, 2017, 2019

© Издательство «Горячая линия – Телеком», 2019



Введение



Цифровая обработка сигналов (ЦОС, DSP — Digital signal processing) стала одной из самых мощных технологий, охвативших науку и технику в XXI веке. Революционные изменения уже коснулись широкого круга областей: коммуникации, медицинской техники, радиолокации и гидролокации, высококачественного воспроизведения музыки, разведки нефтяных месторождений и многих других. В каждой из этих областей произошло глубокое проникновение технологий цифровой обработки сигналов, обеспечившее разработку собственных алгоритмов и специальных методов. Эта комбинация «ширины» и «глубины» проникновения цифровой обработки сигналов делает невозможным индивидуально овладеть всей технологией, которая уже разработана к настоящему времени. Изучение цифровой обработки сигналов включает в себя две задачи: усвоение основных принципов и изучение специализированных методов, которые необходимы для применения ЦОС в конкретных областях науки и техники.

Бурный прогресс вычислительной техники в последние десятилетия привел к широкому внедрению методов цифровой обработки информации практически во всех областях научных исследований и народнохозяйственной деятельности. ЦОС относится к одному из наиболее динамично развивающихся и перспективных направлений и имеет большое фундаментальное и прикладное значение в современной радиотехнике и смежных с нею областях. Ее возможности далеко не исчерпаны как в алгоритмическом, так и в техническом отношении. Это обусловлено созданием новых принципов цифровой обработки и интенсивным развитием элементной базы ЦОС: ростом степени интеграции, функциональной сложности и быстродействия цифровых специализированных процессоров. Методы и алгоритмы цифровой обработки «старые», как методы Ньютона и Гаусса, и «молодые», как компьютеры и специализированные интегральные схемы. Формулы классического численного анализа, такие как формулы для интерполяции, интегрирования и дифференцирования с поправкой на цифровизацию, безусловно, являются алгоритмами цифровой обработки, а наличие быстродействующих цифровых процессоров благоприятствует развитию все более сложных и рациональных алгоритмов обработки сигналов.

ЦОС представляет собой математику, алгоритмы и методы, используемые для обработки сигналов, после того как они были переведены в дискретную форму и, соответственно, выделяется из других областей знаний вычислительной техники уникальным типом данных, который она использует, — цифровые сигналы. История ЦОС начинается в 60-х годах XX века, когда появились первые цифровые компьютеры, но в то время они были дорогими и поэтому область применения ЦОС была весьма ограничена. Революция, произведенная персональными компьютерами в 80-х и 90-х годах XX века, расширила перечень приложений, использующих ЦОС. Кроме военных и правительственных нужд технология цифровой обработки сигналов широко начала применяться в бытовой сфере. ЦОС получила распространение в таких массовых приложениях, как мобильные телефоны, проигрыватели компакт-дисков и электронная речевая почта. Техническая революция привела к тому, что в настоящее время ЦОС является неотъемлемой дисциплиной, включенной в курс обучения для получения базового высшего технического образования. Сегодня технология цифровой обработки сигналов представлена базовыми знаниями, которые необходимы ученым и инженерам.

Курс ЦОС занимает одно из центральных мест среди дисциплин профессиональной подготовки не только специалистов в области цифровой связи, разработчиков и пользователей радиотехнических систем самого различного назначения, но и всех специалистов, в той или иной мере связанных с регистрацией, обращением, обработкой и использованием информационных данных самой различной природы — пользователей цифровых систем. Это определяется тем, что информация, наряду с материей и энергией, принадлежит к фундаментальным философским категориям естествознания и является одной из движущих сил современного развития науки, техники и человеческой цивилизации в целом. Но информация не относится к числу материальных объектов и не существует в явном физическом виде. В ЦОС носителями информации являются цифровые сигналы в любой форме их материального представления в пределах систем, вне которых понятия сигналов также не имеют смысла. Все это приводит к тому, что профессионально грамотная и эффективная регистрация информации, ее обработка, интерпретация и использование возможны только при хорошем знании методов и систем ЦОС.

В настоящее время имеется достаточно большое количество разнообразной литературы по предметной области ЦОС, особенностями данного пособия являются доступность изложения изучаемого материала, значительное количество иллюстраций и максимальная приближенность к структуре изучаемого курса.

Обучение методам и средствам ЦОС осуществляется в рамках дисциплины «Цифровая обработка сигналов». Она охватывает широкий круг теоретических вопросов, изучаемых на лекционных и групповых занятиях, а также в процессе самостоятельной работы обучающихся.

Настоящее учебное пособие представляет собой законченный труд, состоящий из 7 базовых глав, отражающих основную предметную область изучения цифровой обработки сигналов. Структура пособия максимально приближена к тематическому плану изучения дисциплины, а его наполнение отражает необходимый и достаточный теоретический материал в соответствии с программой дисциплины «Цифровая обработка сигналов».

Главы 3, 6 и 7 написаны к.т.н, доцентом А.А. Афанасьевым, разделы 1.1–1.13, 1.15, 1.16, а также главы 2, 3, 4 и разделы 5.3, 5.4.1, 5.5.1–5.5.3, 5.5.5, 5.6, 6.1, 6.3.3–6.3.4 — к.т.н, доцентом А.А. Рыболов-левым, глава 5 — к.т.н. А.П. Рыжковым, остальные разделы подгото-
товлены авторами совместно.



1 Дискретные и цифровые сигналы



1.1. Предмет и проблематика цифровой обработки сигналов

Цифровая обработка сигналов (ЦОС, в английском написании: DSP — digital signal processing), как область науки и техники, включает в себя весьма широкий круг вопросов — от математической теории дискретных и цифровых сигналов и систем до проблем практической реализации устройств. Математические основы ЦОС разработаны достаточно давно, однако эффективная реализация вычислительных алгоритмов долго сдерживалась отсутствием необходимой элементной базы. Значительный прогресс последних десятилетий в области микроэлектроники (повышение степени интеграции и быстродействия интегральных схем, появление и бурное развитие микропроцессоров) обусловил сегодняшнюю высокую динамичность развития практики ЦОС, что выражается в широком распространении цифровых устройств и, в свою очередь, стимулирует дальнейшие теоретические исследования. Основными, тесно взаимосвязанными направлениями современной теории ЦОС можно считать цифровую частотную селекцию сигналов, разработку быстрых алгоритмов обработки сигналов, адаптивную обработку цифровых сигналов, многоскоростную цифровую обработку сигналов и цифровую обработку многомерных сигналов. Широкий спектр решаемых задач, требования выполнения вычислений в реальном масштабе времени привели к созданию новых специальных классов микроэлектронной компонентной базы, получивших название «цифровые процессоры обработки сигналов» (ЦПОС), «программируемые логические интегральные схемы» (ПЛИС) и «системы на кристалле» (СнК). Следует иметь в виду, что для обозначения ЦПОС часто используется другой термин — «сигнальные процессоры». Современные ЦПОС, ПЛИС и СнК, обладая приспособленной для ЦОС архитектурой, позволили значительно повысить эффективность систем об-

работки и передачи сигналов. Теория ЦОС все в большей степени приобретает прикладной характер, при этом упор делается на особенности использования известных и новых алгоритмов и методов общей теории ЦОС в конкретных областях применения с учетом ограничений, накладываемых внутренними ресурсами ЦПОС, ПЛИС и СнК.

Цифровая обработка сигналов изучает методы, алгоритмы и средства обработки дискретных (цифровых) сигналов на основе вычислительного аппарата математики с использованием средств вычислительной техники. Предмет исследований теории ЦОС полностью связан с самим процессом обработки цифровых сигналов в конкретной вычислительной среде и не зависит от цели преобразований, которая определяется областью применения результатов вычислений. Возможность обработки непрерывных сигналов методами ЦОС подразумевает предварительный переход от аналоговых сигналов к сигналам, адекватным аппаратным цифровым средствам. В настоящее время процедуры ЦОС используются в компьютерных технологиях и роботостроении, в устройствах гидро- и радиолокации, в диагностических системах и медицинском оборудовании, в автомобилестроении и авиатехнике, системах управления и т. д. Актуальность изучения проблем ЦОС специалистами в области связи объясняется широким использованием цифровой обработки сигналов в сфере телекоммуникаций. Исторически именно эта сфера явилась одной из базовых для внедрения цифровых технологий и обеспечила условия для их быстрого прогресса. Сегодня совершенно привычными являются цифровая аудио- и видеотехника, цифровая телефония, быстрыми темпами развивается цифровое телевидение, на рынке телекоммуникационного оборудования очевидно доминируют цифровые системы связи. Цифровизация телекоммуникационного поля привела к появлению нового рода услуг, персонализации и доступности разнообразного информационного сервиса.

Тенденция совместного использования телекоммуникационного оборудования и компьютерной техники обусловила появление нового термина — «инфокоммуникации», активно используемого специалистами. Эта тенденция способствует появлению совершенно новых информационно-телекоммуникационных систем, возможности которых по обеспечению абонентов своевременной и качественной информацией еще несколько лет назад казались недостижимыми. Очевидно, что в современных инфокоммуникационных системах доминируют именно цифровые методы хранения, обработки и передачи всех видов информации.

1.2. Функциональная схема системы ЦОС.

Достоинства и недостатки ЦОС

Достаточно типичным для современных инфокоммуникационных систем, на вход которых зачастую поступают непрерывные по природе первичные сигналы (речь, изображение), является случай обработки аналоговых сигналов методами ЦОС. На рис. 1.1 представлена функциональная схема аппаратно реализованной системы цифровой обработки сигналов различных типов.

В качестве частного случая аналого-цифрового преобразования на рис. 1.1 показаны основные процедуры импульсно-кодовой модуляции. Цифровые сигналы $x_{\text{ц}}(nT)$ и $x_{\text{ц}}(nT_1)$ отличаются способом представления (как правило, в десятичном и двоичном представлении), а следовательно, отличаются значением периода дискретизации. В частном случае соотношение периодов T и T_1 определяется количеством двоичных символов в кодовой комбинации кодера. Точки включения в систему источников и получателей сигналов зависят от используемых ими типов сигналов. Так, например, в кодерах современных систем цифровой радиотелефонии ЦПОС реализует алгоритм параметрического сжатия речевого сигнала, а в качестве получателя цифрового сигнала выступает тракт передачи радиотелефонного терминала.

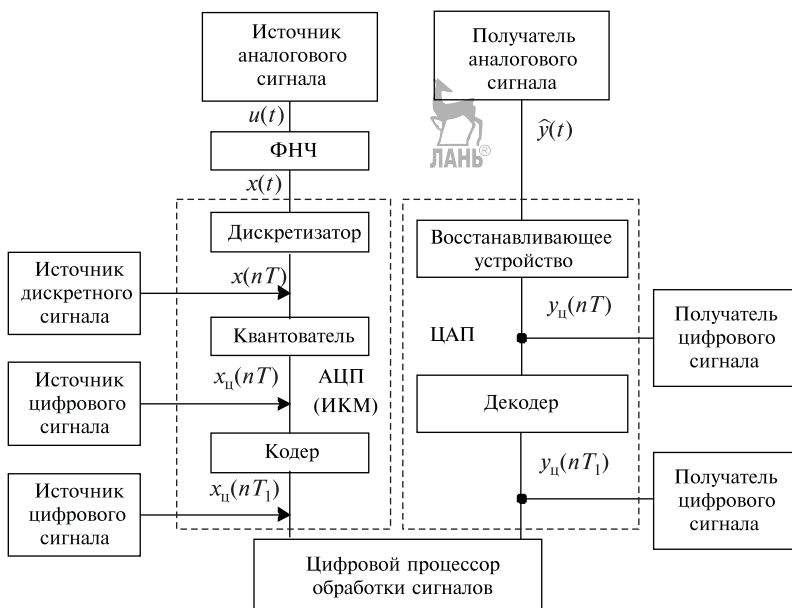


Рис. 1.1. Функциональная схема системы цифровой обработки сигналов

Основные достоинства обработки сигналов цифровыми методами можно классифицировать на три основных вида: принципиальные, реализационные и технико-эксплуатационные.

Среди принципиальных достоинств наиболее значимыми представляются многофункциональность цифровых устройств, возможность мультиплексирования процедур обработки сигналов от нескольких источников, отсутствие принципиальных ограничений на сложность реализуемых алгоритмов. Многофункциональность цифровых устройств проявляется в возможности исполнения одним и тем же цифровым устройством различных алгоритмов в разные моменты времени только за счет изменения используемого программного обеспечения. Возможность мультиплексирования процедур обработки обусловлена тем, что во многих практических случаях частота дискретизации f_d обрабатываемого аналогового сигнала значительно меньше рабочей частоты ЦПОС, в результате чего имеется возможность одновременной реализации одним процессором процедур обработки сигналов от нескольких источников.

К наиболее значимым реализационным достоинствам ЦОС следует отнести стабильность характеристик цифровых устройств, высокую степень их адаптации, точность выполнения вычислительных алгоритмов. Стабильность характеристик устройств ЦОС определяется значительной защитой последних от традиционных дестабилизирующих факторов (температура, влажность и т.д.). В таких условиях верность работы устройств определяется стабильностью частоты задающего генератора и правильностью функционирования арифметико-логических элементов вычислителя. Негативное дестабилизирующее влияние оказывают проникающая радиация и эффекты старения компонентов. Высокая степень адаптации устройств ЦОС объясняется программным характером перестройки алгоритма обработки без необходимости внесения изменений в аппаратную часть устройств.

Основными технико-эксплуатационными достоинствами ЦОС являются высокая надежность, малые массогабаритные показатели цифровых устройств, их малое энергопотребление и широкие диагностические возможности, унификация оборудования и его высокая степень защиты от электромагнитного влияния.

Перечисленные характеристики обеспечивают значительное превосходство устройств ЦОС над аналогичными радиотехническими устройствами, выполненными на основе традиционных технологий. Их совокупность позволяет при сопоставимых стоимостных затратах реализовывать сверхсложные алгоритмы обработки сигналов, которые до самого недавнего времени были лишь предме-

том интереса теоретиков и казались нереализуемыми на практике.

Указанные достоинства ЦОС обуславливают основные преимущества систем цифровой связи по сравнению с системами аналоговой связи:

- легкость восстановления формы цифрового сигнала (так, например, в системах двоичной связи передаваемый цифровой сигнал имеет лишь два возможных состояния);
- высокую надежность цифровых каналов связи;
- более гибкую реализацию систем связи с учетом возможности модернизации и замены программного обеспечения;
- унификацию представления сигналов от источников информации различных типов (речь, изображение, дискретные данные и др.);
- возможность группирования цифровых сигналов в пакеты (ячейки);
- более простую реализацию вспомогательных процедур обработки сигналов (помехоустойчивое кодирование, шифрование и др.).

К основным недостаткам ЦОС следует отнести *относительную ограниченность частотного диапазона* и *нежелательные шумы квантования*, связанные с представлением чисел в устройствах ЦОС и конечной разрядностью используемых компонентов. В целом же главными недостатками цифровых систем связи по сравнению с аналоговыми являются:

- более высокая интенсивность обработки сигналов;
- необходимость выделения значительных ресурсов на многоуровневую синхронизацию систем связи;
- пороговый характер ухудшения качества связи.

Следует отметить, что основным критерием качества цифровых систем связи, в отличие от аналоговых, является вероятность ошибки приема символа (сигнала, цифры).

1.3. Классификация сигналов и их математическое описание

Сигналом называется физический процесс, отображающий сообщение (несущий информацию). Изучение цифровой обработки сигналов требует знания системы классификации сигналов по непрерывности или дискретности их состояний (значений, величин) и особенностей изменения во времени. Известно, что в зависимости от области определения и области изменения временных функций, моделирующих сигналы, различают:

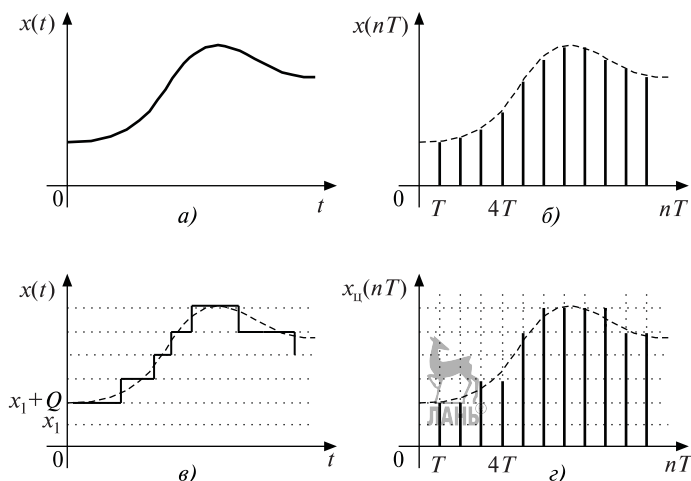


Рис. 1.2. Графическое представление сигналов во временной области

- сигналы, непрерывные по состоянию и времени (аналоговые сигналы) (рис. 1.2,а);
- сигналы, дискретные (квантованные) по состоянию и непрерывные по времени (дискретно-непрерывные сигналы) (рис. 1.2,б);
- сигналы, непрерывные по состоянию и дискретные по времени (дискретные сигналы) (рис. 1.2,в);
- сигналы, дискретные (квантованные) по состоянию и по времени (цифровые сигналы) (рис. 1.2,г).

Сигналы первого вида задаются на конечном или бесконечном временном интервале и могут принимать любые значения в некотором диапазоне. Примерами являются сигналы на выходах микрофона, датчиков температуры, давления и т. д. В качестве моделей аналоговых сигналов выступают функции непрерывного времени $x(t)$.

Дискретно-непрерывные сигналы, часто называемые квантованными по уровню, задаются на некотором временном интервале и характеризуются тем, что принимают только вполне определенные дискретные значения. Такие сигналы могут быть получены из аналоговых сигналов с помощью процедуры квантования по уровню. В результате этой операции непрерывный сигнал заменяется ступенчатой функцией. Шаг квантования Q (расстояние между двумя соседними разрешенными уровнями) может быть как постоянным, так и переменным. Выбор шага квантования определяется с учетом требуемой точности восстановления непрерывного сигнала из квантованного.

Дискретные сигналы задаются в определенные дискретные мо-

менты времени и могут принимать любые значения в некотором диапазоне. Такие сигналы могут быть получены из непрерывных путем взятия отсчетов (выборок, дискретов) в определенные моменты. Это преобразование называется дискретизацией. Шаг дискретизации T (промежуток времени между двумя соседними отсчетами) может быть как переменным, так и постоянным, в последнем случае шаг дискретизации часто называется периодом дискретизации. Выбор T определяется на основе допустимой погрешности при восстановлении аналогового сигнала по последовательности его отсчетов. Дискретные сигналы являются основным объектом теории ЦОС, при этом наиболее часто используется их представление во временной, спектральной и z -областях. Во временной области дискретные сигналы могут быть заданы аналитически, графически или последовательностью отсчетов (например, в виде таблицы), математическими моделями таких сигналов выступают решетчатые функции, обозначаемые $x(nT)$, $x(n)$ или x_n . Аргумент n является безразмерной величиной, представляет собой нормированное время

$$n = \frac{t}{T} = 0, 1, 2, 3, \dots$$

и определяет номер текущего отсчета сигнала. В некоторых случаях аргумент n может принимать и отрицательные значения, при этом следует отметить, что отсчеты незадержанной и физически реализуемой последовательности $x(n)$ могут принимать ненулевые значения только при $n \geq 0$.

Реализация алгоритмов ЦОС цифровыми устройствами требует представления отсчетов сигналов в виде чисел конечной разрядности, что обуславливает переход на практике от дискретных сигналов к цифровым, заданным в определенные дискретные моменты времени и принимающим определенные дискретные значения. Такие сигналы могут быть получены из аналоговых на основе операций дискретизации по времени и квантования по уровню. Математическими моделями цифровых сигналов во временной области выступают квантованные решетчатые функции $x_{\text{ц}}(nT)$, $x_{\text{ц}}(n)$, связанные с решетчатыми функциями дискретных сигналов операторами квантования:

$$x_{\text{ц}}(nT) = F_{\text{кв}}\{x(nT)\}.$$

Конечное число возможных значений решетчатых квантованных функций определяется разрядностью используемых вычислителей и во многом определяет величину искажений, вносимых в аналоговый сигнал при его обработке методами ЦОС.

Очевидно, что все перечисленные виды сигналов могут иметь конечную длину (быть финитными) или быть бесконечными, являться как периодическими, так и непериодическими, вещественными или комплексными и т. д.

1.4. Квантование чисел в ЦОС

Фундаментальным отличием цифрового сигнала от дискретного является использование операции квантования, приводящей (в случае скалярного квантования) к замене исходного числа A квантованным значением $A_{\text{кв}}$. Существенно, что процедуре квантования могут подвергаться как отсчеты обрабатываемого сигнала, так и параметры системы обработки. Необходимость такого подхода обусловлена конечной разрядностью используемых в системах ЦОС операционных устройств (регистров, ячеек памяти, сумматоров, умножителей и т. д.).

Операция квантования является нелинейной. Введение этой операции приводит к неточной реализации алгоритма обработки по отношению к исходному дискретному сигналу. Возникающие при этом ошибки называются ошибками квантования $e = A_{\text{кв}} - A$. Очевидно, что ошибка квантования зависит от шага квантования Q , под которым понимается расстояние между соседними уровнями квантования. В общем случае $Q \neq \text{const}$, однако далее полагается использование равномерного квантования, при этом шаг квантования постоянен и определяется весом младшего разряда кодовой комбинации, представляющей собой квантованное число: $Q = 2^{-b}$, где b — количество числовых разрядов кодового слова.

При квантовании используются два способа: округление и усечение. При округлении исходное число заменяется на ближайшее разрешенное (ближайший уровень квантования), при этом диапазон значений ошибки квантования определяется неравенством $-Q/2 \leq e \leq Q/2$, что означает непревышение ошибкой квантования половины шага квантования. Рабочая характеристика квантователя, использующего способ округления, показана на рис. 1.3,а. На рис. 1.3,б представлена функция плотности распределения вероятности $p(e)$ ошибки квантования округлением; ее равномерность обусловлена допущением о равновероятности всех возможных значений ошибки квантования.

В этих условиях математическое ожидание и дисперсия ошибки квантования e при округлении выражаются следующим образом:

$$\mu_e = 0; \quad \sigma_e^2 = \frac{Q^2}{12} = \frac{2^{-2b}}{12}$$

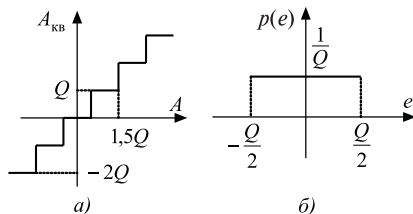


Рис. 1.3. Характеристики равномерного квантования округлением

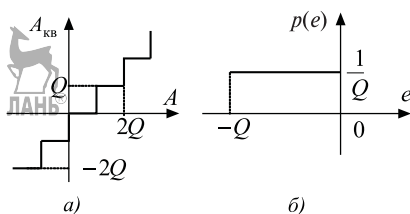


Рис. 1.4. Характеристики равномерного квантования усечением

или

$$\sigma_e^2 = 10 \lg \left(\frac{2^{-2b}}{12} \right) \approx -(6,02b + 10,79) \text{ дБ},$$

где $Q = 2^{-b} = 2^{-m+1}$ — шаг квантования; m — число разрядов в кодовой комбинации.

При усечении исходное число заменяется ближайшим к нему меньшим разрешенным значением (уровнем квантования). Диапазон значений ошибки квантования определяется неравенством $-Q \leq e \leq 0$. Рабочая характеристика квантователя, использующего способ усечения, и соответствующая функция плотности распределения ошибки квантования показаны на рис. 1.4, а и б.

Математическое ожидание и дисперсия ошибки квантования e при усечении выражаются следующим образом:

$$\mu_e = -\frac{Q}{2}; \quad \sigma_e^2 = \frac{Q^2}{12} = \frac{2^{-2b}}{12}$$

или

$$\sigma_e^2 = 10 \lg \left(\frac{2^{-2b}}{12} \right) \approx -(6,02 \cdot b + 10,79) \text{ дБ},$$

где $Q = 2^{-b} = 2^{-m+1}$ — шаг квантования.

1.5. Анализ шумов квантования сигналов

Квантование является нелинейной операцией и вносит в представление квантуемого числа A ошибку $e = A_{\text{кв}} - A$. Существенно, что квантованию в ЦОС могут подвергаться как дискретные (в аналого-цифровых преобразователях), так и цифровые (на выходах операционных узлов процессоров) сигналы; в квантованном виде хранятся и параметры (коэффициенты) системы. Однако квантование коэффициентов системы выполняется однократно, не влияет на линейность системы, а лишь несколько изменяет ее характеристики. Такое изменение прогнозируемо, учитывается на этапе проектирования системы и в дальнейшем материале не рассматривается.

При анализе эффектов квантования в цифровых системах используются как нелинейная, так и линейная модели процедуры квантования, а далее, если это не оговаривается особо, — линейная модель квантования.

Линейная шумовая модель системы ЦОС основана на линейной модели квантования, показанной на рис. 1.5.

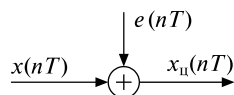


Рис. 1.5. Линейная модель квантования

При линейном моделировании процедуры квантования дискретного сигнала $x(nT)$, представляющего собой последовательность чисел, ошибка квантования выражается дискретным сигналом $e(nT) = x_{\text{ц}}(nT) - x(nT)$, где $x_{\text{ц}}(nT)$ — дискретный квантованный (цифровой) сигнал, связанный с сигналом $x(nT)$ оператором квантования $F_{\text{кв}}$:

$$x_{\text{ц}}(nT) = F_{\text{кв}}[x(nT)].$$

Относительно сигнала $e(nT)$ делается ряд принципиальных допущений:

- последовательность $e(nT)$ является стационарным и эргодическим случайным процессом;
- функция плотности распределения вероятности $p(e)$ ошибки квантования является равномерной в диапазоне возможных значений $e(nT)$;
- любые два отсчета сигнала $e(nT)$ не коррелированы, т. е. сигнал $e(nT)$ представляет собой дискретный случайный процесс типа «белый шум»;
- последовательность $e(nT)$ не коррелирована с квантуемой последовательностью $x(nT)$.

Указанные допущения обуславливают возможность применения для описания сигнала ошибки квантования $e(nT)$ термина «шум квантования» и в значительной степени упрощают проведение анализа эффектов квантования сигналов в системах ЦОС.

На порядок расчета и величину шумов квантования влияют используемый способ квантования (округление или усечение), формат представления чисел, форма представления отсчетов сигнала (с ФТ или с ПТ), используемый код представления значений сигналов (прямой, обратный или дополнительный). В дальнейшем подразумевается, что операции квантования подвергаются числа, представленные в форме с ФТ (с выполнением нормировки $0 \leq |A| < 1$) в дополнительном коде, а при квантовании используются как округление, так и усечение.

Формат представления чисел в системе ЦОС определяется разрядностью функциональных узлов системы. При этом к числу наиболее существенных с точки зрения анализа эффектов квантования относятся:

- m_{Π} — разрядность ячеек памяти;
- $m_{\text{умн}}$ — разрядность выходного регистра умножителя;
- $m_{\text{асс}}$ — разрядность выходного регистра сумматора;
- $m_{\text{вх}}$ — разрядность отсчетов входного сигнала, определяемая, как правило, разрядностью $m_{\text{АЦП}}$ выхода АЦП;
- $m_{\text{вых}}$ — разрядность отсчетов выходного сигнала, определяемая, как правило, разрядностью $m_{\text{ЦАП}}$ входа ЦАП.

Основными точками квантования сигналов в системе ЦОС являются:

- вход системы; при этом формируется шум квантования $e^{\text{АЦП}}(nT)$ входного сигнала;
- выходы операционных узлов вычислителя (как правило — сумматоров); формируемый шум квантования $e^c(nT)$ называется собственным шумом цифрового устройства (системы ЦОС);
- выход с процессора на ЦАП; если в точке выхода используется квантование, то формируется шум квантования $e^{\text{вых}}(nT)$ выходного сигнала.

При этом, как правило, полагается статистическая независимость частных шумовых сигналов, возникающих в этих точках. Следовательно, *полный выходной шум квантования* $e_{\text{вых}}(nT)$ на выходе системы ЦОС, рассчитанный с помощью линейной шумовой модели, может быть определен суммированием указанных шумовых составляющих, пересчитанных к точке выхода системы:

$$e_{\text{вых}}(nT) = e_{\text{вых}}^{\text{АЦП}}(nT) + e_{\text{вых}}^c(nT) + e_{\text{вых}}^{\text{вых}}(nT).$$

Математическое ожидание $\mu_{\text{вых}}$ и дисперсия $(\sigma_{\text{вых}})^2$ полного выходного шума квантования $e_{\text{вых}}(nT)$ определяются в виде сумм математических ожиданий и дисперсий соответствующих составляющих:

$$\begin{aligned} \mu_{\text{вых}} &= \mu_{\text{вых}}^{\text{АЦП}} + \mu_{\text{вых}}^c + \mu_{\text{вых}}^{\text{вых}}, \\ (\sigma_{\text{вых}})^2 &= (\sigma_{\text{вых}}^{\text{АЦП}})^2 + (\sigma_{\text{вых}}^c)^2 + (\sigma_{\text{вых}}^{\text{вых}})^2. \end{aligned}$$

Шумы квантования входного сигнала. На практике часто операция квантования фактически выполняется в периферийных устройствах системы, однако при анализе эффектов квантования ее целесообразно относить ко входу процессора (системы). Отсчеты входного сигнала поступают на вход системы ЦОС (процессора) либо с АЦП (в этом случае $m_{\text{вх}} = m_{\text{АЦП}}$), либо с выхода другой цифровой системы. Очевидно, что в общем случае $m_{\text{вх}} = m_{\text{АЦП}} \leq m_{\Pi}$.

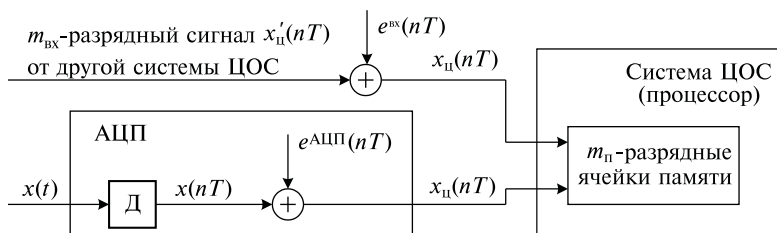


Рис. 1.6. Варианты формирования шумов квантования входного сигнала

Следовательно, возможными причинами шумов квантования входного сигнала (рис. 1.6) являются:

- переход от дискретного сигнала $x(nT)$, формируемого дискретизатором АЦП, к квантованному $m_{\text{АЦП}}$ разрядному сигналу $x_{\text{ц}}(nT)$; при этом возникает шум квантования $e^{\text{АЦП}}(nT)$;
- переход от $m_{\text{вх}}$ разрядного сигнала $x'_{\text{ц}}(nT)$ другой системы ЦОС к $m_{\text{п}}$ -разрядному сигналу $x_{\text{ц}}(nT)$ (в случае $m_{\text{п}} < m_{\text{вх}}$); при этом возникает шум квантования $e^{\text{вх}}(nT)$.

Гипотетически возможна еще одна разновидность входного шума квантования, обусловленная несовпадением разрядности АЦП системы и ячеек памяти процессора ($m_{\text{АЦП}} \neq m_{\text{п}}$), однако, как правило, возможность возникновения такого шума устраняется на этапе выбора элементной базы для реализации системы.

Шумы квантования выходного сигнала. Цифровой $m_{\text{п}}$ -разрядный сигнал $y_{\text{ц}}(nT)$, полученный в результате обработки в системе ЦОС, может в дальнейшем как использоваться в цифровом виде, так и преобразовываться цифро-аналоговым преобразователем в непрерывную форму. В последнем случае при условии $m_{\text{п}} > m_{\text{ЦАП}}$ возникает шум квантования $e^{\text{вых}}(nT)$ выходного сигнала (рис. 1.7) при переходе от разрядности $m_{\text{п}}$ к разрядности $m_{\text{вых}} = m_{\text{ЦАП}}$. Процедура квантования может осуществляться периферийными устройствами, однако для удобства анализа ее следует отнести к точке выхода из системы ЦОС.

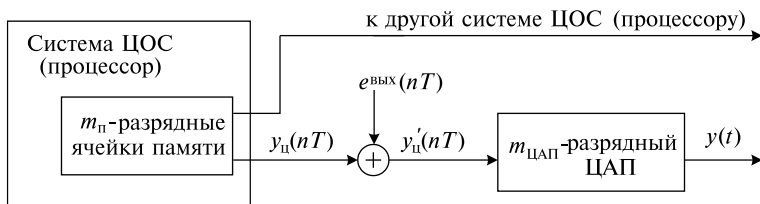


Рис. 1.7. Формирование шума квантования выходного сигнала

1.6. Представление и кодирование чисел в ЦОС

Как правило, числа в ЦОС представляются в двоичной системе счисления, что обусловлено удобством обработки двоичной информации. Достаточно часто для более наглядного компактного отображения двоичной информации используется шестнадцатеричная система счисления.

Представление чисел (данных) в ЦОС обусловлено:

- разрядностью ячеек памяти и регистров;
- условным функциональным распределением разрядов (на знаковые, значащие и т. п.) в ячейках и регистрах;
- спецификой выполнения арифметических операций в процессоре (типом арифметики, используемой в процессоре, и т. д.).

Соответственно представление данных характеризуется форматом, формой и кодом.

Формат чисел связан с разрядностью ячеек памяти и регистров, в которых они хранятся; он определяет возможную длину представления данных в процессоре. Основными форматами представления чисел в процессорах являются:

- слово;
- двойное слово;
- расширенное слово;
- байт;
- полуслово.

Длина слова n измеряется в битах и равна разрядности соответствующей ячейки памяти или регистра; слово обычно характеризует внешнее представление исходных и конечных данных.

Двойное слово отображает содержимое пары соседних n -разрядных ячеек памяти, пары соседних n -разрядных регистров или одного $2n$ -разрядного регистра; двойное слово характеризует представление данных с удвоенной точностью, а его длина равна $2n$ битам.

Расширенное слово отображает содержимое k -разрядного аккумулятора или выходного регистра и характеризует внутреннее представление данных — результатов промежуточных и конечных вычислений. Длина расширенного слова равна k битам и всегда больше длины слова. В зависимости от архитектуры процессора, формы представления данных и назначения регистра для k выполняются следующие условия:

- $k > 2n$ (в процессорах с фиксированной точкой);
- $n < k < 2n$ (в процессорах с плавающей точкой).

Байт и полуслово отображают содержимое соответствующих частей ячейки памяти или регистров.

Очевидно, что формат данных определяет длину последовательности из нулей и единиц — количество битов в двоичной последовательности. Однако для того, чтобы такая последовательность воспринималась как численная величина, необходима дополнительная информация о функциональном распределении битов в последовательности, типе числа, представляемого в заданном формате, и форме представления числа.

Форма представления численных данных (чисел) в процессоре отображает алгебраическую форму записи числа:

- обычной форме записи, например 6000; 6000,0; $-26,34$; 859; 36,571 и т. п., соответствует форма представления чисел с фиксированной точкой (ФТ);
- показательной форме записи, например $6 \cdot 10^3$; $6,0 \cdot 10^3$; $-0,2634 \times 10^2$; $3657,1 \cdot 10^{-2}$ и т. п., соответствует форма представления чисел с плавающей точкой (ПТ).

Примечательно, что в цифровой вычислительной технике принято отделять целую часть числа от дробной не запятой, как в алгебраических формах записи чисел, а точкой.

Формы представления чисел с ФТ и ПТ в различных семействах процессоров могут иметь некоторые особенности, основные закономерности использования этих форм показаны ниже.

Представление чисел в форме с фиксированной точкой.

При представлении числа в форме с ФТ считается, что положение точки, отделяющей целую часть числа от дробной, фиксировано. Разряды слева от точки представляют целую часть числа и его знак, а справа — дробную часть числа. В этом случае каждый разряд числа имеет определенный вес, что позволяет просто реализовывать арифметические операции. Как правило, в системах ЦОС с ФТ используется нормирование обрабатываемых данных таким образом, чтобы все арифметические операции выполнялись с числами $A_{\text{КВ}}$, по абсолютному значению меньшими единицы:

$$0 \leq |A_{\text{КВ}}| < 1. \quad (1.1)$$

В этом случае число содержит $m = b + 1$ двоичных разрядов (рис. 1.8). Старший разряд (наиболее значащий бит, MSF — Most Significant Bit) служит для фиксации знака и называется знаковым (0 — для положительных чисел и нуля, 1 — для отрицательных чисел). Следующие b разрядов служат для фиксации дробной части числа и называются *числовыми*. Точка фиксирована между знаковым и числовыми разрядами. Следовательно, такая форма позволяет представить 2^b отличающихся по абсолютному значению чисел с шагом 2^{-b} (вес наименее значащего бита, LSF — Least Significant

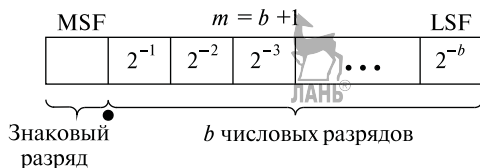


Рис. 1.8. Вариант представления числа в форме с ФТ

Bit) и удовлетворяющих условию (1.1) в диапазоне

$$0 \leq |A_{\text{кв}}| \leq 1 - 2^{-b}. \quad (1.2)$$

Пример 1.1. При использовании для представления чисел, удовлетворяющих условию (1.1), в форме с ФТ восьми двоичных разрядов ($m = 8$) десятичное число 0,5859375 имеет вид 01001011.

В общем случае число, представленное в форме с ФТ и не отвечающее условию (1.1), содержит $m = b_{\text{ц}} + b + 1$ двоичных разрядов (рис. 1.9), причем для представления целой части числа используют $b_{\text{ц}}$ двоичных разрядов, а для представления дробной части числа $-b$ двоичных разрядов.

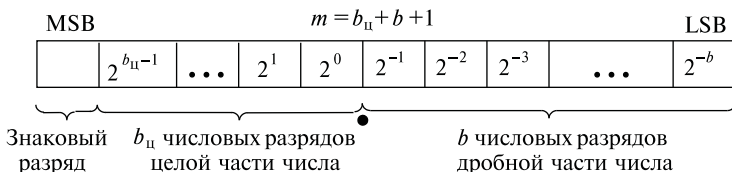


Рис. 1.9. Вариант представления числа в форме с ФТ

В этом случае форма представления чисел с ФТ позволяет представить $2^{b_{\text{ц}}+b}$ чисел, отличающихся по модулю, в диапазоне

$$0 \leq |A_{\text{кв}}| \leq 2^{b_{\text{ц}}} - 2^{-b}. \quad (1.3)$$

В том случае, когда результат арифметической операции выходит за верхний предел выражений (1.2) и (1.3), происходит переполнение регистра, приводящее к искажению результата.

Представление чисел в форме с плавающей точкой. Известно, что в двоичной системе счисления при использовании показательной формы записи каждое вещественное число можно представить бесконечным числом вариантов.

Пример 1.2. Равнозначными являются следующие варианты представления модуля двоичного числа в форме с ФТ:

$$101,011 = 101,011 \cdot 2^0 = 10,1011 \cdot 2^1 = \underline{1,01011} \cdot 2^2 = 1010,11 \cdot 2^{-1} = \dots$$

Для устранения неоднозначности и упрощения арифметики чисел с ПТ из всех возможных вариантов представлений выбран один,

который получил название *нормализованной формы числа с ПТ*. Нормализованная форма числа с ПТ соответствует такому варианту его алгебраической записи, при котором мантисса всегда (за исключением числа нуль) содержит единицу в целой части (в примере этот вариант подчеркнут). Такая нормализация чисел с единицей в целой части мантиссы позволяет при заданном формате увеличить количество значащих цифр вещественного числа на одну, так как бит в целой части мантиссы является неявным и физически не хранится.

С учетом изложенного выше в системах ЦОС форма представления чисел с ПТ принимает вид

$$C = (-1)^S \cdot 2^E \cdot 1, f, \quad (1.4)$$

где C — двоичное число с ПТ; S — знак (0 — положительные числа и нуль, 1 — отрицательные); 2 — основание двоичной системы счисления; E — порядок — целое десятичное число со знаком; $1, f$ — мантисса — вещественное двоичное число без знака, представленное в форме с ФТ; 1 — целая часть мантиссы (неявно присутствующая), f — дробная часть мантиссы.

На представление чисел в форме с ПТ существует единый промышленный стандарт IEEE 754, разработанный в Институте инженеров по электротехнике и электронике (Institute of Electrical and Electronics Engineers) в США в 1985 году. В соответствии с этим стандартом форма представления чисел с ПТ имеет вид

$$C = (-1)^S \cdot 2^e \cdot 1, f,$$

где, в отличие от (1.4), $e = E + \text{bias}$ — смещенный порядок — целое положительное десятичное число; bias — смещение — целая положительная константа. Значение смещения bias выбирается таким образом, чтобы смещенный порядок e всегда имел положительное значение. Это позволяет быстро сравнивать два вещественных числа одинаковых формата и знака.

Представление чисел с ПТ в формате 4-байтового слова в соответствии со стандартом IEEE 754 показано на рис. 1.10.

В этом случае в структуре слова выделяются следующие поля:

- поле знака S — 1 бит;
- поле смещенного порядка e — 8 бит;

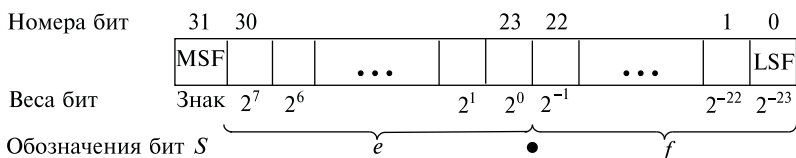


Рис. 1.10. Стандарт IEEE 754 в формате 4-байтового слова

- поле дробной части мантииссы f — 23 бита; целая часть мантииссы, равная 1, присутствует неявно.

Основными характеристиками такого представления являются:

- значение смещения $\text{bias} = 127$;
- $E_{\min} = -126$;
- $E_{\max} = +127$;
- $e_{\min} = +1$;
- $e_{\max} = +254$.



Пример 1.3. Десятичное отрицательное число -180 имеет вид $-180 = (-1)^1 \cdot 2^7 \cdot 1,40625$, из чего следуют равенства $S = 1$, $E = 7$ и 1 , $f = 1,40625$.

При его представлении в формате 4-байтового слова $\text{bias} = 127_{(10)} = 1111111_{(2)}$. Следовательно, смещенный порядок e определяется следующим образом:

$$e = E + \text{bias} = 7_{(10)} + 127_{(10)} = 134_{(10)} = 00000111_{(2)} + 01111111_{(2)} = 10000110_{(2)}.$$

Значение мантииссы:

$$1, f = 1,40625_{(10)} = 2^0 + 2^{-2} + 2^{-3} + 2^{-5} = 1,01101000000000000000_{(2)}.$$

Следует отметить, что диапазон представления чисел в форме с ПТ существенно больше, чем в форме с ФТ. Вероятность переполнения разрядной сетки при выполнении операций с числами с ПТ незначительна. Однако выполнение арифметических операций над числами с ПТ требует большего вычислительного ресурса процессора. В дальнейшем материале данного пособия используется представление чисел в форме с фиксированной точкой.

Кодирование чисел в цифровой обработке сигналов. Ниже представлены прямой, обратный и дополнительный коды, используемые для кодирования положительных и отрицательных двоичных чисел с ФТ $A_{\text{КВ}} = \pm 0.a_1 a_2 a_3 \dots a_b$, удовлетворяющих условию (1.1). Кодовая комбинация в этом случае содержит $m = b + 1$ разрядов.

Прямой код $[A_{\text{КВ}}]_{\text{пр}}$ формируется следующим образом: в знаковый разряд кода записывается 0 для положительных чисел или 1 для отрицательных, числовые разряды кода соответствуют числовым разрядам дробной части числа:

$$[A_{\text{КВ}}]_{\text{пр}} = \begin{cases} 0, a_1 a_2 \dots a_b & \text{при } A_{\text{КВ}} \geq 0; \\ 1, a_1 a_2 \dots a_b & \text{при } A_{\text{КВ}} \leq 0. \end{cases}$$

Условная точка в записи кода отделяет (для наглядности) знаковый разряд от числовых.

Обратный код $[A_{\text{КВ}}]_{\text{обр}}$ при записи положительных чисел соответствует прямому коду. При кодировании отрицательных чисел в знаковый разряд кода записывается 1, а числовые разряды исходного кода инвертируются:

$$[A_{\text{КВ}}]_{\text{обр}} = \begin{cases} 0, a_1 a_2 \dots a_b & \text{при } A_{\text{КВ}} \geq 0, \\ 1, \bar{a}_1 \bar{a}_2 \dots \bar{a}_b & \text{при } A_{\text{КВ}} \leq 0. \end{cases}$$

Дополнительный код $[A_{\text{кв}}]_{\text{доп}}$ при записи положительных чисел также соответствует прямому коду. Правило кодирования отрицательных чисел изменяется: в знаковый разряд кода записывается 1, числовые разряды исходного кода инвертируются (0 заменяется на 1 и наоборот) и к младшему разряду кода добавляется 1 с соблюдением правил сложения двоичных чисел:

$$[A_{\text{кв}}]_{\text{доп}} = \begin{cases} 0, a_1 a_2 \dots a_b & \text{при } A_{\text{кв}} \geq 0; \\ 1, \bar{a}_1 \bar{a}_2 \dots \bar{a}_b + 1 & \text{при } A_{\text{кв}} \leq 0. \end{cases}$$

Исключением является представление в дополнительном коде числа -1 комбинацией $[A_{\text{кв}}]_{\text{доп}} = 1000\dots 0$.

В табл. 1.1 показан пример представления десятичных чисел с помощью 3-разрядных прямого, обратного и дополнительного кодов.

Таблица 1.5

Кодирование чисел с фиксированной точкой

Десятичное число	0,75	0,50	0,25	0		-0,25	-0,50	-0,75	-1,00
Прямой код	011	010	001	000	100	101	110	111	—
Обратный код	011	010	001	000	111	110	101	100	—
Дополнительный код	011	010	001	000		111	110	101	100

Особенностью дополнительного кода является однозначность представления числа 0. Кроме того, преимуществом дополнительного кода является то, что все арифметические операции с числами независимо от их типа (целые или дробные) выполняются одинаково как с беззнаковыми числами, когда знаковый разряд воспринимается как старший значащий. Результат арифметической операции в этом случае тоже представляется в дополнительном коде.

Ниже рассмотрен пример сложения в дополнительном коде чисел, представленных в формате байта в форме с фиксированной точкой, предварительно прошедших нормировку (1.1).

Пример 1.4. Пусть в десятичном представлении пример имеет вид $-0,8125 + 0,6484375 = -0,1640625$.

Формы записи слагаемых многовариантны:

$$\begin{aligned} -0,8125_{(10)} &= 11101000_{(\text{пр})} = 10010111_{(\text{обр})} = 10011000_{(\text{доп})}; \\ +0,6484375_{(10)} &= 01010011_{(\text{пр})} = 01010011_{(\text{обр})} = 01010011_{(\text{доп})}. \end{aligned}$$

Для дополнительного кода рассматриваемый пример выражается следующей записью:

$$10011000 + 01010011 \neq 11101011.$$

Нетрудно убедиться, что комбинация 11101011 в дополнительном коде представляет число $-0,1640625$:

$$-0,1640625_{(10)} = 10010101_{(\text{пр})} = 11101010_{(\text{обр})} = 11101011_{(\text{доп})}.$$

Таким образом, число -180 в стандарте IEEE 754 в формате 4-байтового слова имеет вид 11000011001101000000000000000000.

Так как в ЦОС наиболее часто используется дополнительный код, в дальнейшем материале, если иное не оговорено особо, используется представление чисел в форме с фиксированной точкой в дополнительном коде.

Таким образом, особенности процедур дискретизации, квантования и кодирования в значительной степени характеризуют конкретную систему цифровой обработки сигналов и влияют на ее рабочие характеристики.

1.7. Типовые дискретные сигналы

Ниже представлены основные типовые дискретные сигналы, достаточно часто используемые в теории и практике ЦОС.

Цифровой единичный импульс $\delta(nT)$ определяется следующей последовательностью:

$$\delta(nT) = \begin{cases} 1, & n = 0; \\ 0, & n \neq 0. \end{cases}$$

т. е. $\delta(nT) = \{1, 0, 0, 0, \dots\}$. В дискретных системах цифровой единичный импульс играет такую же роль, как δ -функция в аналоговых системах, но, в отличие от последней, является физически реализуемым сигналом. На рис. 1.11, а и б представлены сигналы $\delta(nT)$ и $\delta(nT - n_0T)$.

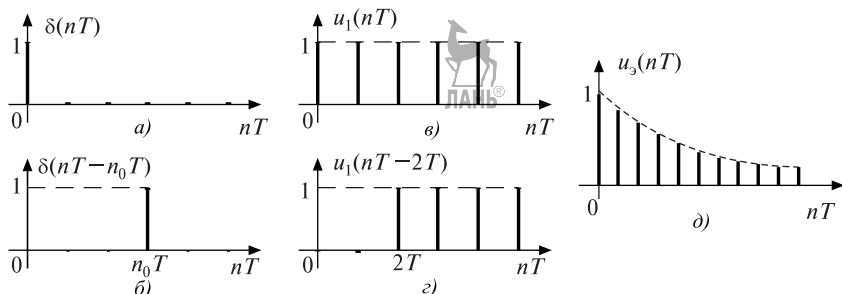


Рис. 1.11. Графическое представление типовых дискретных сигналов во временной области

В ряде случаев произвольный дискретный сигнал $x(nT)$ удобно представлять с помощью цифрового единичного импульса в виде

$$x(nT) = \sum_{k=0}^{\infty} x(kT)\delta(nT - kT), \quad n = 0, 1, 2, 3, \dots$$

Пример 1.5. Конечный дискретный сигнал $x(n) = \{2, 1, 3\}$ может быть представлен с использованием цифрового единичного импульса в виде

$$x(n) = \sum_{k=0}^2 x(k)\delta(n-k), \quad n = 0, 1, 2;$$

$$x(0) = x(0)\delta(0-0) + x(1)\delta(0-1) + x(2)\delta(0-2) = x(0)\delta(0-0) = 2;$$

$$x(1) = x(0)\delta(1-0) + x(1)\delta(1-1) + x(2)\delta(1-2) = x(1)\delta(1-1) = 1;$$

$$x(2) = x(0)\delta(2-0) + x(1)\delta(2-1) + x(2)\delta(2-2) = x(2)\delta(2-2) = 3.$$

Цифровой единичный скачок $u_1(nT)$ определяется как

$$u_1(nT) = \begin{cases} 1, & n \geq 0; \\ 0, & n < 0, \end{cases}$$

и является дискретным аналогом единичной ступенчатой функции (функции Хевисайда), используемой в теории непрерывных сигналов.

На рис. 1.11, в и г представлены сигнал $u_1(nT)$ и задержанный на два периода дискретизации цифровой единичный скачок $u_1(nT - 2T) = \{0, 0, 1, 1, 1, \dots\}$ соответственно.

Дискретная экспонента определяется выражением

$$u_a(nT) = \begin{cases} a^n, & n \geq 0; \\ 0, & n < 0. \end{cases}$$

При $0 < a < 1$ дискретная экспонента является монотонно убывающей функцией (рис. 1.11, д), при $a > 1$ — монотонно возрастающей, а в случае $a = 1$ превращается в цифровой единичный скачок.

Дискретная косинусоида соответствует выражению $u_c(nT) = A \cos(\omega nT) = A \cos(2\pi f nT)$ и может быть получена дискретизацией аналогового гармонического колебания.

Пример 1.6. Аналоговый гармонический сигнал $x(t) = A \cos(2\pi ft)$ имеет частоту $f = 0,5$ кГц. Определить дискретную косинусоиду, получаемую за счет дискретизации исходного колебания с частотой $f_d = 4$ кГц.

Получаемый сигнал имеет вид

$$u_c(nT) = A \cos(2\pi f nT) = A \cos(2\pi f_d n) = A \cos(0,25\pi n).$$

Безразмерная величина

$$f_0 = fT = \frac{f}{f_d}$$

носит название нормированной частоты и достаточно часто используется при дискретных представлениях сигналов. Очевидно, что в рассмотренном примере $f_0 = 0,125$. На рис. 1.12 графически представлены сигналы $x(t)$, $u_c(nT)$ и $u_c(n)$ в соответствии с примером 1.6.

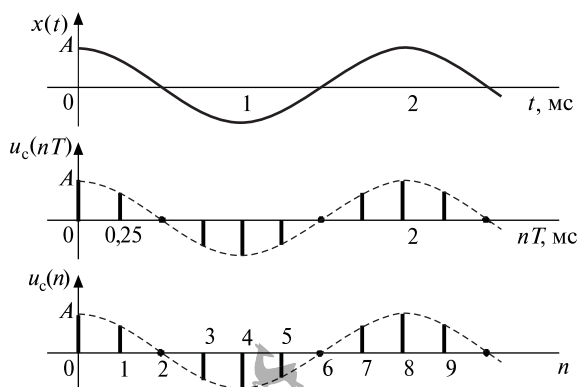


Рис. 1.12. Графическое представление дискретной косинусоиды

1.8. Основные операции цифровой обработки сигналов

Простейшими операциями, осуществляемыми над дискретными сигналами, являются сдвиг сигнала, алгебраическое сложение сигналов, их перемножение и умножение сигнала на константу. Более сложной операцией, часто используемой в ЦОС, является свертка дискретных сигналов. Рассмотрим представление этих операций во временной области.

Сдвиг дискретного сигнала. Сдвиг исходной последовательности $x(nT)$ на m отсчетов вправо (при $m > 0$) или влево (при $m < 0$) приводит к образованию сдвинутой последовательности $y(nT) = x(nT - mT)$.

Пример 1.7. На рис. 1.13 изображены исходная последовательность $x(nT) = \{1, 3, 2, 0, 2\}$ (рис. 1.13,а) и последовательности $y_1(nT) = x(nT - 2T)$ (рис. 1.13,б) и $y_2(nT) = x(nT + 2T)$ (рис. 1.13,в), сдвинутые относительно исходной.

На практике операция сдвига реализуется линией задержки на mT . Например, для варианта рис. 1.13,б линия задержки изображена на рис. 1.14.

Алгебраическое сложение дискретных сигналов. Алгебраическое сложение сигналов $x_1(nT)$ и $x_2(nT)$ заключается в формировании дискретной последовательности $y(nT)$, дискреты которой представляют собой суммы соответствующих дискретов исходных последовательностей: $y(nT) = x_1(nT) + x_2(nT)$.

Пример 1.8. Пусть: $x_1(nT) = \{3, 0, 2, 3, 1\}$, $x_2(nT) = \{0, 1, 1, 1, 0\}$. Тогда $y(nT) = x_1(nT) + x_2(nT) = \{3, 1, 3, 4, 1\}$.

На практике операция алгебраического сложения реализуется сумматором (рис. 1.15).

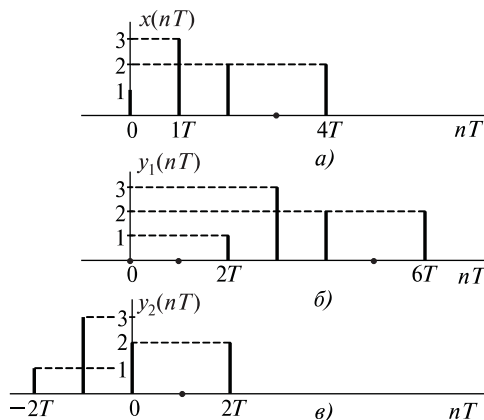


Рис. 1.13. Графическое представление сдвига дискретного сигнала

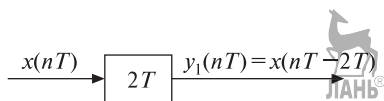
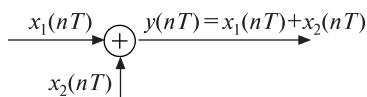
Рис. 1.14. Реализация задержки дискретного сигнала на $2T$ 

Рис. 1.15. Реализация сложения дискретных сигналов

Умножение дискретного сигнала на константу. Умножение дискретного сигнала $x(nT)$ на константу a заключается в формировании дискретной последовательности $y(nT)$, дискреты которой представляют собой произведения соответствующих дискретов исходной последовательности на эту константу: $y(nT) = ax(nT)$.

Пример 1.9. Пусть $x(nT) = \{2, 1, 3, 1\}$, $a = 2$. Тогда $y(nT) = ax(nT) = \{2 \cdot 2, 2 \cdot 1, 2 \cdot 3, 2 \cdot 1\} = \{4, 2, 6, 2\}$.

На практике операция реализуется умножителем (рис. 1.16).

Перемножение дискретных сигналов. Перемножение дискретных сигналов $x_1(nT)$ и $x_2(nT)$ заключается в формировании дискретной последовательности $y(nT)$, дискреты которой представляют собой произведения соответствующих дискретов исходных последовательностей: $y(nT) = x_1(nT)x_2(nT)$.

Пример 1.10. Пусть $x_1(nT) = \{1, 2, 0, 3\}$, $x_2(nT) = \{1, 2, 3, 1\}$. Тогда $y(nT) = x_1(nT)x_2(nT) = \{1 \cdot 1, 2 \cdot 2, 0 \cdot 3, 3 \cdot 1\} = \{1, 4, 0, 3\}$.

Практическая реализация перемножения дискретных сигналов показана на рис. 1.17.

Свертка дискретных сигналов. Сверткой сигналов $x_1(nT)$ и $x_2(nT)$ называется операция, определяемая выражением

$$y(nT) = \sum_{m=0}^{\infty} x_1(mT)x_2(nT - mT) = \sum_{m=0}^{\infty} x_1(nT - mT)x_2(mT),$$

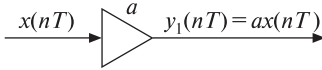


Рис. 1.16. Реализация умножения дискретного сигнала на константу

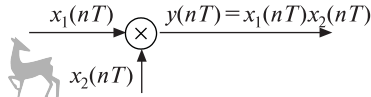


Рис. 1.17. Реализация перемножения дискретных сигналов

при этом выбор варианта расчета определяется удобством вычисления в конкретном случае. В большом числе практических применений свертки верхний предел суммирования может быть заменен на n . Одним из традиционных применений свертки дискретных сигналов является расчет реакции дискретной системы (например, цифрового фильтра) на воздействие, в этом случае один дискретный сигнал соответствует входному сигналу системы, а второй — ее импульсной характеристике.

Частными вариантами свертки дискретных сигналов являются *периодическая* (круговая) и *апериодическая* (линейная) свертки.

Периодическая (круговая) свертка определена для периодических последовательностей $x_1(nT)$ и $x_2(nT)$, имеющих одинаковые периоды NT . Последовательность $y(nT)$, представляющая периодическую свертку сигналов $x_1(nT)$ и $x_2(nT)$, также имеет период в N отсчетов и определяется следующим образом:

$$y(nT) = \sum_{m=0}^{N-1} x_1(mT)x_2(nT - mT) = \sum_{m=0}^{N-1} x_1(nT - mT)x_2(mT),$$

при этом значения сигналов $x_1(nT)$ и $x_2(nT)$ за пределами периода следует при расчете считать нулевыми.

Пример 1.11. Пусть $x_1(nT) = \{2, 1, 3\}$, $x_2(nT) = \{1, -1, 0\}$. Тогда

$$y(0) = x_1(0)x_2(0) + x_1(T)x_2(-T) + x_1(2T)x_2(-2T) = x_1(0)x_2(0) = 2;$$

$$y(T) = x_1(0)x_2(T) + x_1(T)x_2(0) + x_1(2T)x_2(-T) = 2(-1) + 1 \cdot 1 = -1;$$

$$y(2T) = x_1(0)x_2(2T) + x_1(T)x_2(T) + x_1(2T)x_2(0) = 2 \cdot 0 + 1 \cdot (-1) + 3 \cdot 1 = 2.$$

Следовательно, $y(nT) = \{2, -1, 2\}$.

Апериодическая (линейная) свертка определена для конечных дискретных последовательностей $x_1(nT)$ длиной в N_1 отсчетов и $x_2(nT)$ длиной в N_2 отсчетов, она имеет конечное число отсчетов $N = N_1 + N_2 - 1$ и определяется следующим образом:

$$y(nT) = \sum_{m=0}^n x_1(mT)x_2(nT - mT) = \sum_{m=0}^n x_1(nT - mT)x_2(mT).$$

Пример 1.12. Пусть $x_1(nT) = \{1, 2\}$, $x_2(nT) = \{-2, 1, 2\}$. Тогда

$$y(0) = x_1(0)x_2(0) = 1 \cdot (-2) = -2;$$

$$y(T) = x_1(0)x_2(T) + x_1(T)x_2(0) = 1 \cdot 1 + 2 \cdot (-2) = -3;$$

$$y(2T) = x_1(0)x_2(2T) + x_1(T)x_2(T) + x_1(2T)x_2(0) = 1 \cdot 2 + 2 \cdot 1 = 4;$$

$$y(3T) = x_1(0)x_2(3T) + x_1(T)x_2(2T) + x_1(2T)x_2(T) + x_1(3T)x_2(0) = 2 \cdot 2 = 4.$$

Следовательно, $y(nT) = \{-2, -3, 4, 4\}$.

1.9. Спектры непериодических дискретных сигналов и их свойства

Во многих задачах ЦОС целесообразно использовать спектральную форму представления дискретных сигналов, т. е. представлять сигналы в виде совокупности элементарных функций. В теории обработки радиотехнических сигналов с этой целью применяется большое число преобразований, использующих различные системы базисных функций. Наибольшее распространение исторически получили разновидности преобразований Фурье для различных классов сигналов.

Спектром $X(e^{j\omega T})$ дискретного сигнала $x(nT)$ называется прямое преобразование Фурье вида

$$X(e^{j\omega T}) = \sum_{n=-\infty}^{\infty} x(nT) \exp(-j\omega nT). \quad (1.5)$$

Выражение (1.5) позволяет перейти от временного представления сигнала к его представлению в спектральной области. Обратный переход во временную область осуществляется на основе обратного преобразования Фурье для дискретного сигнала:

$$x(nT) = \frac{T}{2\pi} \int_{-\pi/T}^{\pi/T} X(e^{j\omega T}) \exp(j\omega nT) d\omega. \quad (1.6)$$

Ниже представлены свойства спектра дискретного сигнала.

1. Анализ выражения (1.5) позволяет сделать вывод о том, что спектр дискретного сигнала является периодической по частоте функцией с периодом, равным частоте дискретизации $\omega_d = 2\pi f_d = 2\pi/T$. Модуль и аргумент спектра также являются периодическими функциями с тем же периодом.

2. Модуль спектра (спектр амплитуд) вещественного сигнала $x(nT)$ является четной функцией частоты:

$$|X(e^{j\omega T})| = |X(e^{-j\omega T})|,$$

а аргумент спектра (спектр фаз) — нечетной функцией частоты:

$$\arg X(e^{j\omega T}) = -\arg X(e^{-j\omega T}).$$

3. Вещественная часть спектра вещественного сигнала $x(nT)$ является четной функцией частоты:

$$\operatorname{Re}\{X(e^{j\omega T})\} = \operatorname{Re}\{X(e^{-j\omega T})\},$$

а его мнимая часть — нечетной функцией частоты:

$$\operatorname{Im}\{X(e^{j\omega T})\} = -\operatorname{Im}\{X(e^{-j\omega T})\}.$$

4. Спектр линейной комбинации дискретных сигналов $y(nT) = a_1x_1(nT) + a_2x_2(nT)$ равен такой же линейной комбинации спектров исходных сигналов:

$$Y(e^{j\omega T}) = a_1X_1(e^{j\omega T}) + a_2X_2(e^{j\omega T}).$$

Пример 1.13. Определить спектр непериодического дискретного сигнала $x(nT) = e^{\alpha nT}$, $n = 0, 1, 2, \dots$, $T = 1$ мс, $\alpha = -100$.

Исследуемый сигнал представляет собой бесконечную дискретную последовательность $x(nT) = e^{-100n \cdot 10^{-3}} = e^{-0,1n}$. Отсчеты сигнала при $n = 0, 1, 2, 3, 4$ имеют следующие значения:

$$x(0) = e^0 = 1; \quad x(1T) = e^{-0,1} \approx 0,905; \quad x(2T) = e^{-0,2} \approx 0,819;$$

$$x(3T) = e^{-0,3} \approx 0,74; \quad x(4T) = e^{-0,4} \approx 0,67.$$

Графическое представление исследуемого сигнала на временном интервале $[0; 7T]$ показано на рис. 1.18.

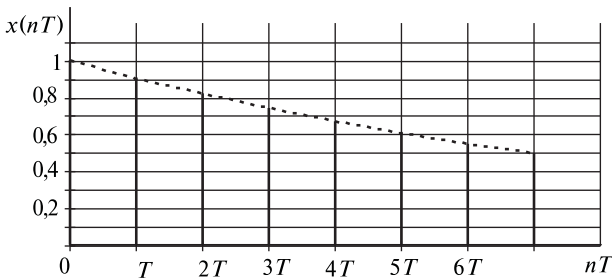


Рис. 1.18. Графическое представление сигнала $x(nT) = e^{-0,1n}$ во временной области

В соответствии с выражением (1.5) спектр исследуемого дискретного непериодического сигнала определяется как

$$X(e^{j\omega T}) = \sum_{n=0}^{\infty} x(nT)e^{-j\omega nT} = \sum_{n=0}^{\infty} e^{\alpha nT} e^{-j\omega nT} = \sum_{n=0}^{\infty} e^{(\alpha - j\omega)nT}.$$

Члены полученного ряда представляют собой бесконечную геометрическую прогрессию, причем первый член отличен от нуля: $a = e^{(\alpha - j\omega)0T} = 1$. Известно, что такой ряд сходится в случае, если знаменатель прогрессии q по абсолютному значению меньше единицы: $|q| < 1$; при этом сумма ряда вычисляется как $S_n = a/(1 - q)$.

Следовательно, при $|q| = |e^{(\alpha - j\omega)T}| < 1$ спектр сигнала $x(nT)$ определяется как

$$X(e^{j\omega T}) = \frac{1}{1 - e^{(\alpha - j\omega)T}} = \frac{1}{1 - e^{\alpha T} e^{-j\omega T}} =$$

$$= \frac{1}{1 - e^{\alpha T}(\cos \omega T - j \sin \omega T)} = \frac{1}{1 - e^{\alpha T} \cos \omega T + j e^{\alpha T} \sin \omega T}.$$

Амплитудный спектр сигнала $x(nT)$, определяемый модулем $|X(e^{j\omega T})|$, может быть представлен в различных частотных шкалах:

$$\text{а) } |X(e^{j\omega T})| = \frac{1}{\sqrt{(1 - e^{\alpha T} \cos \omega T)^2 + (e^{\alpha T} \sin \omega T)^2}} =$$

$$= \frac{1}{\sqrt{[1 - e^{-0,1} \cos(\omega \cdot 0,001)]^2 + [e^{-0,1} \sin(\omega \cdot 0,001)]^2}};$$

$$\text{б) } |X(e^{j2\pi f T})| =$$

$$= \frac{1}{\sqrt{[1 - e^{-0,1} \cos(2\pi f \cdot 0,001)]^2 + [e^{-0,1} \sin(2\pi f \cdot 0,001)]^2}};$$

$$\text{в) } |X(e^{j\omega_0})| = \frac{1}{\sqrt{(1 - e^{-0,1} \cos \omega_0)^2 + (e^{-0,1} \sin \omega_0)^2}};$$

$$\text{г) } |X(e^{j2\pi f_0})| = \frac{1}{\sqrt{(1 - e^{-0,1} \cos 2\pi f_0)^2 + (e^{-0,1} \sin 2\pi f_0)^2}}.$$

Анализ полученных выражений свидетельствует о том, что амплитудный спектр вещественного сигнала (рис. 1.19) $x(nT) = e^{-0,1n}$ является периодической четной функцией частоты с периодом, равным частоте дискретизации сигнала: $\omega_d = 2\pi f_d = 2\pi/T = 2\pi \times 10^3$ рад/с.

Фазовый спектр исследуемого сигнала определяется аргументом комплексного спектра $X(e^{j\omega T})$, является периодической нечет-

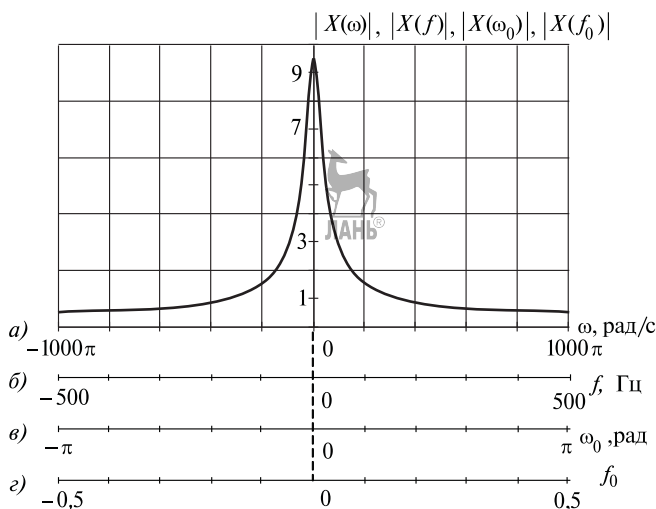


Рис. 1.19. Амплитудный спектр сигнала $x(nT) = e^{-0,1n}$

ной функцией частоты и для различных частотных шкал определяется следующими выражениями:

$$\begin{aligned} \text{а) } \arg\{X(e^{j\omega T})\} &= -\operatorname{arctg} \frac{e^{\alpha T} \sin \omega T}{1 - e^{\alpha T} \cos \omega T} = \\ &= -\operatorname{arctg} \frac{e^{-0,1} \sin(\omega \cdot 0,001)}{1 - e^{-0,1} \cos(\omega \cdot 0,001)}; \end{aligned}$$

$$\text{б) } \arg\{X(e^{j2\pi f T})\} = -\operatorname{arctg} \frac{e^{-0,1} \sin(2\pi f \cdot 0,001)}{1 - e^{-0,1} \cos(2\pi f \cdot 0,001)};$$

$$\text{в) } \arg\{X(e^{j\omega_0})\} = -\operatorname{arctg} \frac{e^{-0,1} \sin \omega_0}{1 - e^{-0,1} \cos \omega_0};$$

$$\text{г) } \arg\{X(e^{j2\pi f_0})\} = -\operatorname{arctg} \frac{e^{-0,1} \sin 2\pi f_0}{1 - e^{-0,1} \cos 2\pi f_0}.$$

Графики полученных выражений фазового спектра на длительности одного периода для соответствующих частотных шкал изображены на рис. 1.20.

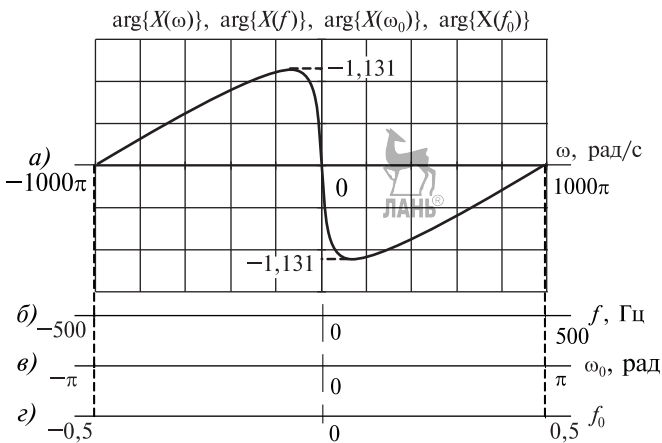


Рис. 1.20. Фазовый спектр сигнала $x(nT) = e^{-0,1n}$

1.10. Соотношение спектров непериодических аналоговых и дискретных сигналов

В практике построения систем связи и управления на основе ЦОС типовыми являются задачи преобразования аналогового сигнала $x(t)$ в дискретный $x(nT)$ и обратно. В связи с этим достаточно важным представляется характер взаимодействия спектра $X(e^{j\omega T})$ дискретной последовательности $x(nT)$ со спектром $X(j\omega)$ исходного аналогового сигнала. Пусть дискретный сигнал $x(nT)$ получен

из аналогового $x(t)$ дискретизацией последнего с периодом T . Пара преобразований Фурье для аналогового сигнала имеет вид:

$$\begin{aligned} X(j\omega) &= \int_{-\infty}^{\infty} x(t)e^{-j\omega t} dt; \\ x(t) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} X(j\omega)e^{j\omega t} d\omega. \end{aligned} \quad (1.7)$$

Для выявления связи между спектрами $X(j\omega)$ и $X(e^{j\omega T})$ воспользуемся равенством $x(nT) = x(t)|_{t=nT}$ и представим второе выражение (1.7) в новом виде, заметив интеграл с бесконечными пределами по угловой частоте ω бесконечной суммой интегралов на частотных интервалах, равных частоте дискретизации $\omega_d = 2\pi/T = 2\pi f_d$:

$$x(nT) = \frac{1}{2\pi} \sum_{m=-\infty}^{\infty} \int_{(2m-1)\pi/T}^{(2m+1)\pi/T} X(j\omega)e^{j\omega nT} d\omega. \quad (1.8)$$

Произведем в выражении (1.8) некоторые изменения:

- изменим очередность суммирования и интегрирования;
- заменим «скользящие» пределы интегрирования $[(2m-1)\pi/T; (2m+1)\pi/T]$ фиксированными $[-\pi/T; \pi/T]$ при соответствующем переходе к «скользящему» значению частоты;
- умножим подынтегральное выражение на коэффициент $1/T$.

Тогда получим

$$x(nT) = \frac{T}{2\pi} \int_{-\pi/T}^{\pi/T} \frac{1}{T} \sum_{m=-\infty}^{\infty} X \left[j \left(\omega + \frac{2\pi}{T} m \right) \right] e^{j\omega nT} d\omega. \quad (1.9)$$

Сравнивая формулы (1.6) и (1.9), получим искомое соотношение спектров аналогового и дискретного сигналов:

$$X(e^{j\omega T}) = \frac{1}{T} \sum_{m=-\infty}^{\infty} X \left[j \left(\omega + \frac{2\pi}{T} m \right) \right] = \sum_{m=-\infty}^{\infty} \frac{1}{T} X[j(\omega + m\omega_d)]. \quad (1.10)$$

Из полученного выражения видно, что спектр дискретного сигнала $X(e^{j\omega T})$, являясь периодической функцией частоты, состоит из бесконечного числа слагаемых. Каждое из слагаемых с номером m представляет собой спектр аналогового сигнала $X(j\omega)$, масштабированный коэффициентом $1/T$ и сдвинутый по частоте на величину $m\omega_d = m\frac{2\pi}{T}$.

Графическая иллюстрация полученного соотношения (1.10) представлена на рис. 1.21. Пусть амплитудный спектр $|X(j\omega)|$ аналогового сигнала $x(t)$ ограничен частотным интервалом $[-\omega_1, \omega_1]$ (рис. 1.21,а). В соответствии с выражением (1.10) спектр $X(e^{j\omega T})$

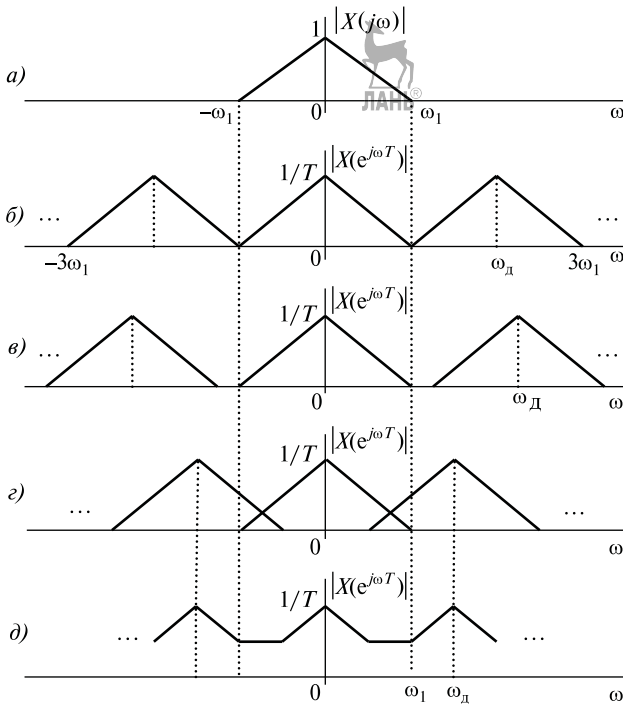


Рис. 1.21. Графики амплитудных спектров аналогового и дискретного сигналов

дискретного сигнала $x(nT)$, полученного дискретизацией сигнала $x(t)$ с частотой дискретизации ω_d , будет составлять

$$X(e^{j\omega T}) = \frac{1}{T} \{ \dots + X[j(\omega - \omega_d)] + X(j\omega) + X[j(\omega + \omega_d)] + \dots \}. \quad (1.11)$$

В случае, если частота дискретизации равна удвоенной максимальной частоте спектра исходного непрерывного сигнала ($\omega_d = 2\omega_1$), слагаемые правой части выражения (1.11) соприкасаются по частотной оси. Амплитудный спектр дискретного сигнала при этом имеет вид, показанный на рис. 1.21, б.

При превышении частотой дискретизации удвоенной максимальной частоты спектра исходного сигнала ($\omega_d > 2\omega_1$) слагаемые правой части выражения (1.11) разнесены по частоте (рис. 1.21, в). В обоих рассмотренных случаях отсутствует явление наложения спектров, при этом в составе спектра дискретного сигнала в неискаженном виде сохраняется спектр исходного аналогового сигнала, что создает условия для последующего безыскаженного восстановления сигнала $x(t)$.

По своей сути полученный результат подтверждает известную

теорему В.А. Котельникова, в соответствии с которой любой непрерывный сигнал со спектром, сосредоточенным в ограниченной полосе частот $[-\omega_1, \omega_1]$, может быть представлен последовательностью своих отсчетов, взятых с частотой $\omega_d \geq 2\omega_1$.

Принципиально иная картина наблюдается в случае $\omega_d < 2\omega_1$: частотное перекрытие слагаемых правой части выражения (1.11) (рис. 1.21,з) приводит к суммарному спектру дискретного сигнала (рис. 1.21,д), не содержащему спектр исходного аналогового сигнала $x(t)$, что делает невозможным безыскаженное восстановление.

Важным следствием проведенного анализа является следующее положение: при любой частоте дискретизации аналоговых сигналов с бесконечным спектром слагаемые правой части выражения (1.11) будут перекрываться по частоте, что неизбежно приведет к искажению восстановленного аналогового сигнала. Следовательно, цифровая обработка непрерывного сигнала с нефинитным спектром в некоторой мере теряет смысл. По этой причине на практике осуществляется предварительное ограничение спектра сигнала $x(t)$ низкочастотным фильтром с частотой среза ω_1 , после чего последующая дискретизация сигнала с частотой $\omega_d \geq 2\omega_1$ приводит к корректной цифровой обработке.

Следует обратить внимание на особенности дискретизации узкополосных аналоговых сигналов, спектр которых сосредоточен в окрестностях некоторой частоты ω_0 и имеет ширину $2\omega_1$ (рис. 1.22).

В соответствии с рис. 1.21 требуемая частота дискретизации такого сигнала составляет $\omega_d \geq 2(\omega_0 + \omega_1)$. Однако анализ, аналогичный проведенному выше, позволяет утверждать, что дискретный сигнал сохраняет всю информацию об исходном узкополосном аналоговом сигнале при дискретизации последнего с частотой $\omega_d \geq 4\omega_1$.

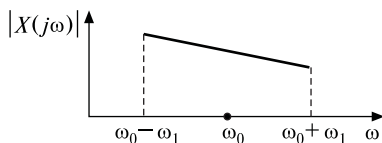


Рис. 1.22. Односторонний амплитудный спектр узкополосного аналогового сигнала

1.11. Спектры периодических дискретных сигналов

Дискретное преобразование Фурье и его свойства. Для спектрального представления периодических дискретных сигналов с периодом в N отсчетов применяются дискретные спектральные преобразования. При этом используются системы таких дискретных базисных функций, у которых моменты отсчетов совпадают с моментами отсчетов исследуемых сигналов. Среди различных дискретных спектральных преобразований в современной теории и практике

ЦОС наиболее широкое применение находит дискретное преобразование Фурье (ДПФ).

В случае ДПФ в состав базисной системы входят ортогональные дискретные экспоненциальные функции вида

$$\varphi_k(nT) = e^{jk\Omega nT} = e^{j(2\pi/N)kn} = W_N^{-kn};$$

где n — порядковый номер отсчетов дискретного сигнала, $n = 0, 1, 2, \dots, N-1$; k — порядковый номер спектральных составляющих дискретного сигнала, $k = 0, 1, 2, \dots, N-1$; $W_N = e^{-j2\pi/N}$ — поворачивающий множитель ДПФ; $\Omega = 2\pi/(NT)$ — основная частота ДПФ («бин» ДПФ).

Пусть дискретный сигнал $x(nT)$ является периодическим:

$$x(nT) = x[(n + mN)T], \quad m = 0, 1, 2, \dots$$

Прямое ДПФ (ПДПФ) реализует переход от временной формы представления периодического дискретного сигнала к спектральной и выражается в виде

$$X(jk\Omega) = \sum_{n=0}^{N-1} x(nT)e^{-j\frac{2\pi}{N}kn} = \sum_{n=0}^{N-1} x(nT)W_N^{kn} = \sum_{n=0}^{N-1} x(nT)e^{-jk\Omega nT}. \quad (1.12)$$

Периодичность (с периодом N) показательных функций в выражении (1.12) свидетельствует о том, что последовательность $X(jk\Omega)$ так же, как и последовательность $x(nT)$, является периодической с периодом в N отсчетов:

$$X(jk\Omega) = X[j(k + lN)\Omega], \quad l = 0, 1, 2, \dots$$

Иначе говоря, комплексный спектр $X(jk\Omega)$, его модуль и аргумент являются периодическими функциями частоты с периодом, равным частоте дискретизации $\omega_d = N\Omega$. Кроме того, для вещественного сигнала $x(nT)$ амплитудный спектр $|X(jk\Omega)|$ представляет собой четную функцию частоты, а фазовый спектр $\arg X(jk\Omega)$ — нечетную.

Противоположную задачу перехода из спектральной области во временную решает обратное ДПФ (ОДПФ):

$$\begin{aligned} x(nT) &= \frac{1}{N} \sum_{k=0}^{N-1} X(jk\Omega)e^{j\frac{2\pi}{N}kn} = \frac{1}{N} \sum_{k=0}^{N-1} X(jk\Omega)W_N^{-kn} = \\ &= \frac{1}{N} \sum_{k=0}^{N-1} X(jk\Omega)e^{jk\Omega nT}. \end{aligned} \quad (1.13)$$

Смысл формулы (1.13) можно пояснить следующим образом: периодический сигнал $x(nT)$, имеющий на периоде N отсчетов, мож-

но с точностью до масштабирующего коэффициента $1/N$ представить в виде суммы N соответствующих комплексных дискретных экспонент, действительные части которых представляют собой косинусоиды с амплитудами $|X(0)|, |X(j \cdot 1\Omega)|, |X(j \cdot 2\Omega)|, \dots, |X[j(N - 1)\Omega]|$, угловыми частотами $0, \Omega, 2\Omega, \dots, (N - 1)\Omega$ и начальными фазами $\arg X(0), \arg X(j1\Omega), \dots, \arg X[j(N - 1)\Omega]$.

ДПФ может применяться не только к периодическим дискретным сигналам, но и к дискретным сигналам конечной длительности, содержащим N отсчетов. Известно, что такие сигналы обладают непрерывными периодическими спектрами. В этом случае результат ДПФ представляет собой дискретные отсчеты непрерывного спектра на периоде, по которым в соответствии с теоремой Котельникова в частотной области можно точно восстановить непрерывный спектр исследуемого финитного дискретного сигнала.

Пример 1.14. Дискретный периодический сигнал $x(nT) = \{1, 1, 1, 0, 0, 0\}$ имеет период дискретизации $T = 1$ мс.

Требуется определить спектр сигнала, восстановить дискретную последовательность по спектральным составляющим.

Графическое представление исследуемого сигнала на временном интервале $[0, 7 \text{ мс}]$ показано на рис. 1.23.

Очевидно, что период исследуемого сигнала содержит $N = 6$ отсчетов и равен 6 мс. Спектр дискретного периодического сигнала определяется на основе прямого ДПФ:

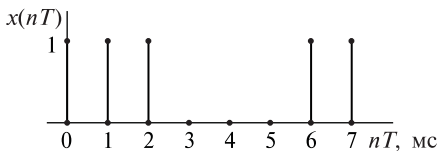


Рис. 1.23. Графическое представление исследуемого сигнала во временной области

$$X(jk\Omega) = \sum_{n=0}^{N-1} x(nT)e^{-j(2\pi/N)kn}, \quad k = 0, 1, 2, \dots, N - 1.$$

Основная угловая частота ДПФ составляет $\Omega = \frac{2\pi}{NT} = \frac{10^3\pi}{3}$ рад/с. Так как при $n = 3, 4, 5$ $x(nT) = 0$, то

$$X(jk\Omega) = \sum_{n=0}^2 x(nT)e^{-j(2\pi/6)kn}, \quad k = 0, 1, 2, \dots, N - 1.$$

Следовательно, отсчеты спектра исследуемого сигнала имеют следующие значения:

$$\begin{aligned} X(j \cdot 0\Omega) &= x(0)e^{-j(2\pi/6) \cdot 0 \cdot 0} + x(T)e^{-j(2\pi/6) \cdot 0 \cdot 1} + x(2T)e^{-j(2\pi/6) \cdot 0 \cdot 2} = \\ &= 1 + 1 + 1 = 3 = 3 + j \cdot 0 = 3e^{j \cdot 0}; \end{aligned}$$

$$\begin{aligned} X(j \cdot 1\Omega) &= x(0)e^{-j(2\pi/6) \cdot 1 \cdot 0} + x(T)e^{-j(2\pi/6) \cdot 1 \cdot 1} + x(2T)e^{-j(2\pi/6) \cdot 1 \cdot 2} = \\ &= 1 + e^{-j\pi/3} + e^{-j2\pi/3} = 1 + \cos \frac{\pi}{3} - j \sin \frac{\pi}{3} + \cos \frac{2\pi}{3} - j \sin \frac{2\pi}{3} = \\ &= 1 + \frac{1}{2} - j \frac{\sqrt{3}}{2} - \frac{1}{2} - j \frac{\sqrt{3}}{2} = 1 - j\sqrt{3} = 2e^{-j\frac{\pi}{3}}; \end{aligned}$$

$$\begin{aligned} X(j \cdot 2\Omega) &= x(0)e^{-j(2\pi/6) \cdot 2 \cdot 0} + x(T)e^{-j(2\pi/6) \cdot 2 \cdot 1} + x(2T)e^{-j(2\pi/6) \cdot 2 \cdot 2} = \\ &= 1 + e^{-j2\pi/3} + e^{-j4\pi/3} = 1 + \cos \frac{2\pi}{3} - j \sin \frac{2\pi}{3} + \cos \frac{4\pi}{3} - j \sin \frac{4\pi}{3} = \end{aligned}$$

$$\begin{aligned}
&= 1 + (-0,5) - j \frac{\sqrt{3}}{2} + (-0,5) + j \frac{\sqrt{3}}{2} = 0; \\
X(j \cdot 3\Omega) &= x(0)e^{-j(2\pi/6) \cdot 3 \cdot 0} + x(T)e^{-j(2\pi/6) \cdot 3 \cdot 1} + x(2T)e^{-j(2\pi/6) \cdot 3 \cdot 2} = \\
&= 1 + e^{-j\pi} + e^{-j2\pi} = 1 + \cos \pi - j \sin \pi + \cos 2\pi - j \sin 2\pi = \\
&= 1 + (-1) - j \cdot 0 + 1 - j \cdot 0 = 1 - j \cdot 0 = 1e^{j0}; \\
X(j \cdot 4\Omega) &= x(0)e^{-j(2\pi/6) \cdot 4 \cdot 0} + x(T)e^{-j(2\pi/6) \cdot 4 \cdot 1} + x(2T)e^{-j(2\pi/6) \cdot 4 \cdot 2} = \\
&= 1 + e^{-j4\pi/3} + e^{-j8\pi/3} = 1 + \cos \frac{4\pi}{3} - j \sin \frac{4\pi}{3} + \cos \frac{8\pi}{3} - j \sin \frac{8\pi}{3} = \\
&= 1 + (-0,5) - j \left(-\frac{\sqrt{3}}{2} \right) + (-0,5) - j \frac{\sqrt{3}}{2} = 0; \\
X(j \cdot 5\Omega) &= x(0)e^{-j(2\pi/6) \cdot 5 \cdot 0} + x(T)e^{-j(2\pi/6) \cdot 5 \cdot 1} + x(2T)e^{-j(2\pi/6) \cdot 5 \cdot 2} = \\
&= 1 + e^{-j5\pi/3} + e^{-j10\pi/3} = 1 + \cos \frac{5\pi}{3} - j \sin \frac{5\pi}{3} + \cos \frac{10\pi}{3} - j \sin \frac{10\pi}{3} = \\
&= 1 + 0,5 - j \left(-\frac{\sqrt{3}}{2} \right) + (-0,5) - j \left(\frac{\sqrt{3}}{2} \right) = 1 + j\sqrt{3} = 2e^{j\pi/3}.
\end{aligned}$$

Основные параметры спектральных составляющих исследуемого сигнала, полученные по результатам расчетов, представлены в табл. 1.2, а амплитудный и фазовый спектры на длительности одного периода показаны на рис. 1.24 и 1.25 соответственно.

Таблица 1.2

Основные параметры спектральных составляющих исследуемого сигнала

k	Амплитуда $A(k\Omega) = X(jk\Omega) $	Начальная фаза $\varphi(k\Omega) = \arg X(jk\Omega)$, рад	Угловая частота $\omega_k = k\Omega$, рад/с	Частота $f_k = \frac{\omega_k}{2\pi}$, Гц	Период $T_k = \frac{1}{f_k}$, мс
0	3	0	0	0	0
1	2	$-\frac{\pi}{3}$	$\frac{10^3 \pi}{3}$	$\frac{10^3}{6}$	6
2	0	0	$\frac{2 \cdot 10^3 \cdot \pi}{3}$	$\frac{10^3}{3}$	3
3	1	0	$10^3 \pi$	$\frac{10^3}{2}$	2
4	0	0	$\frac{4 \cdot 10^3 \pi}{3}$	$\frac{2 \cdot 10^3}{3}$	1,5
5	2	$\frac{\pi}{3}$	$\frac{5 \cdot 10^3 \pi}{3}$	$\frac{5 \cdot 10^3}{6}$	1,2

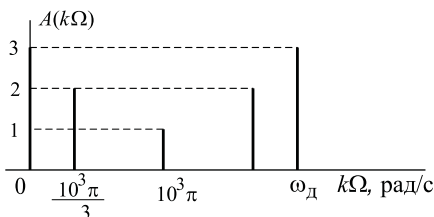


Рис. 1.24. Спектр амплитуд исследуемого сигнала

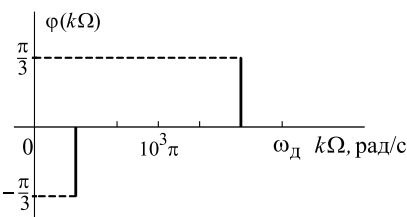


Рис. 1.25. Спектр фаз исследуемого сигнала

Очевидно, что спектр $X(jk\Omega)$ является периодической функцией частоты с периодом $\omega_d = 6\Omega = 2\pi \cdot 10^3$ рад/с.

Следует отметить, что спектр амплитуд является четной функцией частоты, а спектр фаз — нечетной.

Задача восстановления дискретной последовательности по известному спектру сигнала осуществляется процедурой ОДПФ и заключается в синтезе сигнала $x(nT)$ посредством суммирования масштабированных спектральных составляющих:

$$x(nT) = \frac{1}{N} \sum_{k=0}^{N-1} X(jk\Omega) e^{j(2\pi/N)kn}, \quad n = 0, 1, 2, \dots, N-1.$$

Тогда

$$\begin{aligned} x(nT) &= \frac{1}{6} \left[3e^{j(2\pi/6) \cdot 0n} + (1 - j\sqrt{3})e^{j(2\pi/6) \cdot 1n} + 0 + 1e^{j(2\pi/6) \cdot 3n} + 0 + (1 + j\sqrt{3}) \times \right. \\ &\quad \left. \times e^{j(2\pi/6) \cdot 5n} \right] = \frac{1}{6} \left[3 + 2e^{j \arctg(-\sqrt{3})} e^{j(2\pi/6)n} + e^{j\pi n} + 2e^{j \arctg(\sqrt{3})} e^{j(2\pi/6) \cdot 5n} \right] = \\ &= \frac{1}{6} \left[3 + 2e^{-j\pi/3} e^{j(2\pi/6)n} + e^{j\pi n} + 2e^{j\pi/3} e^{j(2\pi/6) \cdot 5n} \right] = \\ &= \frac{1}{6} \left[3 + 2e^{j((2\pi/6)n - \pi/3)} + e^{j\pi n} + 2e^{j((2\pi/6) \cdot 5n + \pi/3)} \right]. \end{aligned}$$

Следовательно, отсчеты восстанавливаемого сигнала $x(nT)$ имеют следующие значения:

$$\begin{aligned} x(0) &= \frac{1}{6} [3 + 2e^{j(-\pi/3)} + 1 + 2e^{j\pi/3}] = \frac{1}{6} (3 + 2 + 1) = 1; \\ x(T) &= \frac{1}{6} [3 + 2e^{j(\pi/3 - \pi/3)} + e^{j\pi} + 2e^{j(10\pi/6 + 2\pi/6)}] = \frac{1}{6} (3 + 2 - 1 + 2) = 1; \\ x(2T) &= \frac{1}{6} [3 + 2e^{j(4\pi/6 - \pi/3)} + e^{j2\pi} + 2e^{j(20\pi/6 + \pi/3)}] = \frac{1}{6} (3 + 2 + 1) = 1; \\ x(3T) &= \frac{1}{6} [3 + 2e^{j(\pi - \pi/3)} + e^{j3\pi} + 2e^{j(5\pi + \pi/3)}] = \frac{1}{6} (3 - 2 - 1) = 0; \\ x(4T) &= \frac{1}{6} [3 + 2e^{j(4\pi/3 - \pi/3)} + e^{j4\pi} + 2e^{j(20\pi/3 + \pi/3)}] = \frac{1}{6} (3 - 2 + 1 - 2) = 0; \\ x(5T) &= \frac{1}{6} [3 + 2e^{j(5\pi/3 - \pi/3)} + e^{j5\pi} + 2e^{j(25\pi/3 + \pi/3)}] = \\ &= \frac{1}{6} [3 + 2e^{j4\pi/3 - 1} + 2e^{j26\pi/3}] = \frac{1}{6} (3 - 2 - 1) = 0. \end{aligned}$$

Полученные значения отсчетов полностью соответствуют отсчетам исходного сигнала, что свидетельствует об однозначности и обратимости процедуры ДПФ. Графическая иллюстрация соответствия исходного сигнала $x(nT)$ и отсчетов дискретных гармоник на длительности одного периода $x(nT)$ представлена на рис. 1.26.

Аналитические выражения дискретных гармоник $x_k(nT)$, являющихся действительными частями $X(jk\Omega)$, в соответствии с обозначениями в табл. 1.2 имеют вид

$$x_k(nT) = \operatorname{Re}[X(jk\Omega)] = A(k\Omega) \cos[k\Omega nT + \varphi(k\Omega)].$$

Очевидно, что непрерывные аналоги этих гармоник, показанные на рис. 1.26 пунктирными линиями, соответствуют выражению

$$x_k(t) = A(k\Omega) \cos[k\Omega t + \varphi(k\Omega)].$$

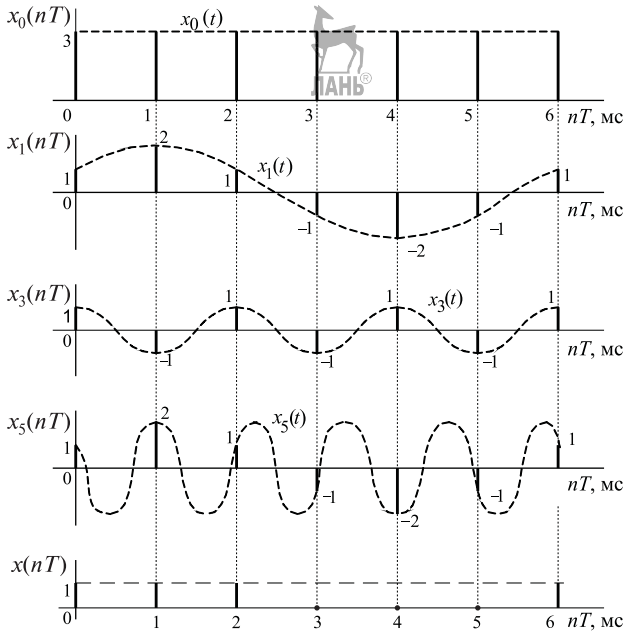


Рис. 1.26. Графическая иллюстрация синтеза $x(nT)$ на основе ОДПФ

Из рис. 1.26 хорошо видно, что $x(nT) = \frac{1}{N} \sum_{k=0}^5 x_k(nT)$, причем $x_2(nT) = x_4(nT) = 0$ (см. табл. 1.2).

Основными свойствами ДПФ являются следующие:

Свойство линейности. Спектр линейной комбинации дискретных сигналов равен такой же линейной комбинации спектров этих сигналов.

Если сигнал $x_1(nT)$ имеет спектр $X_1(jk\Omega)$, $n = 0, 1, 2, \dots, N-1$; $k = 0, 1, 2, \dots, N-1$; а сигнал $x_2(nT)$ спектр $X_2(jk\Omega)$, то сигнал $x_3(nT) = a_1 x_1(nT) + a_2 x_2(nT)$, где a_1 и a_2 — постоянные коэффициенты, имеет спектр $X_3(jk\Omega) = a_1 X_1(jk\Omega) + a_2 X_2(jk\Omega)$.

Свойство спектра периодической свертки. Спектр дискретного сигнала, равного периодической свертке двух исходных сигналов, равен произведению спектров исходных сигналов.

Если сигнал $x_1(nT)$ имеет спектр $X_1(jk\Omega)$, $n, k = 0, 1, 2, \dots, N-1$, а сигнал $x_2(nT)$ — спектр $X_2(jk\Omega)$, то сигнал

$$y(nT) = \sum_{m=0}^{N-1} x_1(mT)x_2(nT - mT) = \sum_{m=0}^{N-1} x_1(nT - mT)x_2(mT)$$

имеет спектр $X_1(jk\Omega)X_2(jk\Omega)$.

Свойство задержки (сдвига). Спектр дискретного сигнала, сдвинутого относительно исходного на m отсчетов, равен произведению спектра исходного сигнала на множитель W_N^{km} .

Если сигнал $x(nT)$ имеет спектр $X(jk\Omega)$, $n, k = 0, 1, 2, \dots, N-1$, то сигнал $x[(n-m)T]$ имеет спектр $X(jk\Omega)W_N^{km}$, $m = 1, 2, \dots, N-1$.

Ниже доказательство данного свойства выполнено в общем виде, а иллюстрация приведена для конкретных исходных данных.

Спектр $X(jk\Omega)$ дискретного периодического сигнала с периодом в N отсчетов определяется дискретным преобразованием Фурье:

$$X(jk\Omega) = \sum_{n=0}^{N-1} x(n) e^{-j \frac{2\pi}{N} kn} = \sum_{n=0}^{N-1} x(n) W_N^{kn}.$$

На рис. 1.27,а представлен вариант сигнала $x(n)$, имеющего период $N = 8$.

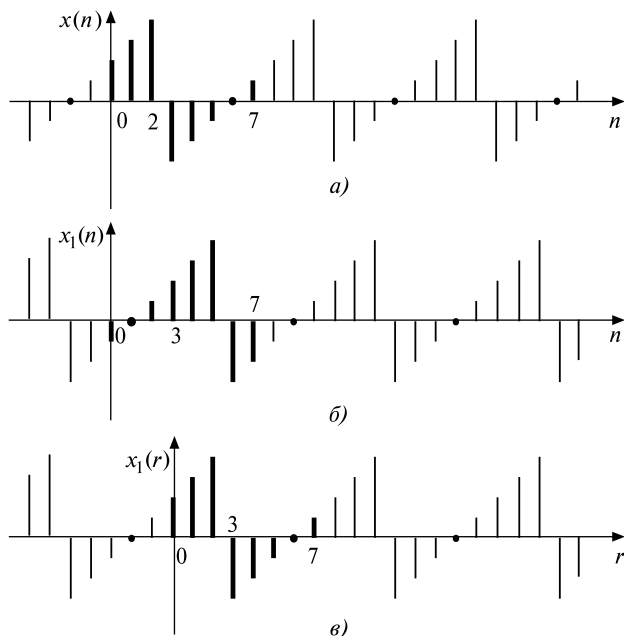


Рис. 1.27. Иллюстрация к расчету спектра дискретного периодического сигнала, имеющего задержку во времени

Дискретный сигнал $x_1(n) = x(n-m)$ является результатом задержки исходного сигнала $x(n)$ на m периодов дискретизации, $m = 1, 2, \dots, N-1$. На рис. 1.27,б показан сигнал $x_1(n) = x(n-3)$.

Существенно, что задержка периодического сигнала не нарушает периодичности задержанного сигнала (вследствие этого такой

сдвиг часто называют циклическим), а следовательно, спектр сигнала $x_1(n)$ также определяется дискретным преобразованием Фурье:

$$X_1(jk\Omega) = \sum_{n=0}^{N-1} x_1(n) e^{-j(2\pi/N)kn} = \sum_{n=0}^{N-1} x(n-m) W_N^{kn}.$$

Для выполнения расчета проведем следующую замену переменной: $n-m=r$, из чего следует $n=r+m$. На рис. 1.27,а представлен сигнал $x_1(r)$. Очевидно, что момент времени $r=0$ соответствует моменту времени $n=3$.

С учетом выполненной замены полученное выражение приобретает следующий вид:

$$X_1(jk\Omega) = \sum_{r=0}^{N-1} x(r) W_N^{k(r+m)} = \sum_{r=0}^{N-1} x(r) W_N^{kr} W_N^{km}.$$

Очевидно, что $\sum_{r=0}^{N-1} x(r) W_N^{kr} = X(jk\Omega)$ (данное равенство достаточно хорошо иллюстрируется сравнением рис. 1.27,а и в). Следовательно, выражение для расчета спектра задержанного сигнала может быть преобразовано к искомому виду:

$$X_1(jk\Omega) = X(jk\Omega) W_N^{km} = X(jk\Omega) e^{-j(2\pi/N)km} = X(jk\Omega) e^{-jk\Omega mT}.$$

Анализ полученного выражения свидетельствует о том, что при задержке сигнала во времени на m периодов дискретизации его амплитудный спектр не меняется, а фазовый спектр получает линейное приращение $-\frac{2\pi}{N}km = -\frac{2\pi}{NT}Tkm = -k\Omega mT$, что означает задержку всех спектральных составляющих сигнала $x(nT)$ на время mT .

Значительный интерес представляет зависимость между ДПФ и рядом Фурье, графическая иллюстрация которой представлена на рис. 1.28.

Непрерывный периодический сигнал $x(t)$ с периодом T_a (рис. 1.28,а) имеет дискретный амплитудный спектр $|X(jk\Omega_a)|$ с максимальной частотой $\omega_{\max} = 4\Omega_a$, изображенный для нормированной частотной шкалы ω/Ω_a на рис. 1.28,в. Дискретный периодический сигнал с периодом T_a (рис. 1.28,б) получен дискретизацией сигнала $x(t)$ с частотой $\omega_d = 2\omega_{\max}$, причем $T_a = 8T$, т.е. $N = 8$. Амплитудный спектр сигнала $x(nT)$ представляет собой модуль прямого ДПФ этого сигнала и является периодической функцией частоты с периодом $\omega_d = 8\Omega$. Один период спектра амплитуд $|X(jk\Omega)|$ для нормированной частотной шкалы ω/Ω изображен на рис. 1.28,г. Из представленных графиков очевидно, что спектральные линии $|X(jk\Omega_a)|$,

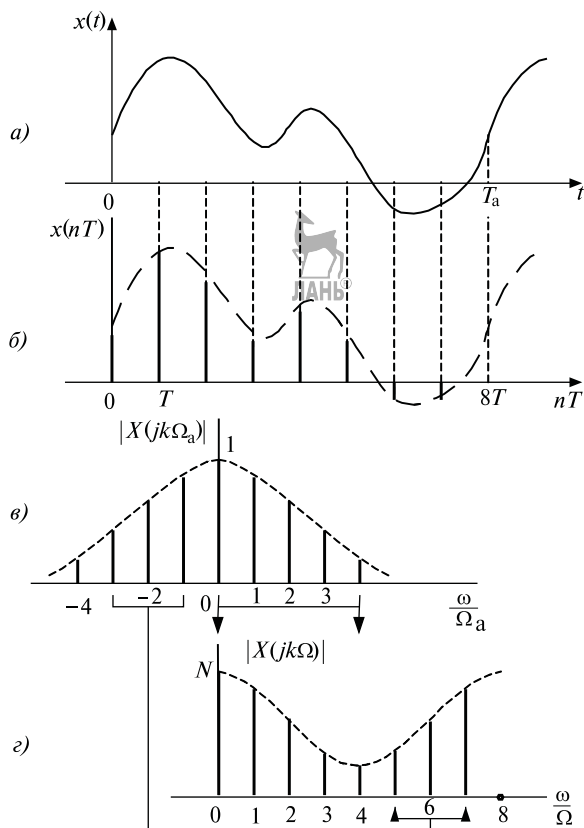


Рис. 1.28. Графическая иллюстрация зависимости между ДПФ и рядом Фурье

расположенные в положительной области частот, полностью соответствуют аналогичным спектральным линиям в спектре $|X(jk\Omega)|$, а спектральные линии $|X(jk\Omega_a)|$, расположенные в отрицательной области частот, продолжают амплитудный спектр $|X(jk\Omega)|$ так, что

$$|X(jk\Omega)| = \begin{cases} N|X(jk\Omega_a)|, & k = 0, 1, 2, \dots, N/2; \\ N \left| X \left[j \left(-\frac{N}{2} + l \right) \Omega_a \right] \right|, & k = \frac{N}{2} + l; l = 1, 2, \dots, \frac{N}{2} - 1. \end{cases}$$

1.12. Характеристика спектрального представления различных классов сигналов на основе преобразований Фурье

Ниже представлена характеристика спектрального представления четырех различных классов сигналов на основе преобразований Фурье по признакам:



- система базисных функций;
- математические выражения для прямого и обратного преобразований Фурье;
- вид спектра.

I. *Непрерывные непериодические сигналы $x(t)$, удовлетворяющие условиям Дирихле.*

1) Система непрерывных по времени функций вида

$$\{\varphi(t)\} = \{\exp(j\omega t)\}, \quad -\infty < \omega < \infty.$$

2) Преобразования Фурье:

$$X(j\omega) = \int_{-\infty}^{\infty} x(t) \exp(-j\omega t) dt; \quad x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} X(j\omega) \exp(j\omega t) d\omega.$$

3) Спектр непрерывный.

II. *Непрерывные периодические сигналы $x(t)$ с периодом T_a , удовлетворяющие условиям Дирихле.*

1) Система непрерывных по времени функций вида

$$\{\varphi_k(t)\} = \{\exp(jk\Omega_a t)\}, \quad \Omega_a = \frac{2\pi}{T_a}; \quad k = \dots, -2, -1, 0, 1, 2, \dots$$

2) Преобразования Фурье:

$$X(jk\Omega_a) = \frac{1}{T_a} \int_{-T_a/2}^{T_a/2} x(t) \exp(-jk\Omega_a t) dt;$$

$$x(t) = \sum_{k=-\infty}^{+\infty} X(jk\Omega_a) \exp(jk\Omega_a t).$$

3) Спектр дискретный с шагом дискретизации Ω_a .

III. *Дискретные непериодические сигналы $x(nT)$ (T — период дискретизации).*

1) Система дискретных по времени функций:

$$\{\varphi(nT)\} = \{\exp(j\omega nT)\}, \quad -\infty < \omega < \infty, \quad n = 0, 1, 2, 3, \dots$$

2) Преобразования Фурье:

$$X(j\omega T) = \sum_{n=-\infty}^{\infty} x(nT) \exp(-j\omega nT);$$

$$x(nT) = \frac{T}{2\pi} \int_{-\pi/T}^{\pi/T} X(j\omega T) \exp(j\omega nT) d\omega.$$

3) Спектр непрерывный периодический с периодом $\omega_d = 2\pi/T$.

IV. *Дискретные периодические сигналы $x(nT)$ с периодом NT .*

1) Система дискретных по времени функций:

$$\{\varphi_k(nT)\} = \{\exp(jk\Omega nT)\}, \quad \Omega = \frac{2\pi}{NT}, \quad n, k = 0, 1, 2, \dots, N-1.$$

2) Дискретное преобразование Фурье:

$$X(jk\Omega) = \sum_{n=0}^{N-1} x(nT) \exp(-jk\Omega nT);$$

$$x(nT) = \frac{1}{N} \sum_{k=0}^{N-1} X(jk\Omega) \exp(jk\Omega nT).$$

3) Спектр дискретный с шагом дискретизации Ω и периодический с периодом $\omega_d = N\Omega = 2\pi/T$.

1.13. Быстрое преобразование Фурье

Для оценки вычислительной сложности дискретного преобразования Фурье представим его в традиционной форме:

$$X(jk\Omega) = \sum_{n=0}^{N-1} x(nT) e^{-jk\Omega nT} = \sum_{n=0}^{N-1} x(nT) W_N^{kn}. \quad (1.14)$$

Непосредственный расчет спектра по формуле (1.14) требует достаточно большого числа вычислений. Действительно, процедура определения k -й частотной составляющей

$$X(jk\Omega) = \sum_{n=0}^{N-1} x(nT) e^{-jk\Omega nT} = x(0) \cdot e^0 \pm x(T) \cdot e^{-jk\Omega T} \pm \pm x(2T) \cdot e^{-jk\Omega 2T} \pm \pm \dots \pm x(NT-T) \cdot e^{-jk\Omega (NT-T)}$$

включает в себя N операций умножения с комплексными числами (подчеркнуты одной чертой) и $N-1$ операций сложения с комплексными числами (подчеркнуты двумя чертами). Так как общее число частотных составляющих равно N , то полная процедура ДПФ требует выполнения N^2 операций умножения и $N(N-1)$ операций сложения. Очевидно, что реализация такого числа операций в реальных схемах, где используются последовательности с $N > 256$, затрудняет реализацию вычислительных алгоритмов в режиме реального времени.

Для уменьшения вычислительной сложности процедуры ДПФ разработаны более рациональные по сравнению с выражением (1.14) алгоритмы, получившие название *быстрые преобразования Фурье* (БПФ, FFT — Fast Fourier Transform). В основе этих алгоритмов лежит процедура деления N -точечной последовательности $x(n)$ на две

более короткие $x_1(n)$ и $x_2(n)$ длительностью, как правило, $N/2$ точек (отсчетов). Очевидно, что расчет ДПФ $X_1(jk\Omega)$ и $X_2(jk\Omega)$ для вновь сформированных последовательностей потребует $(N/2)^2 \cdot 2 = N^2/2$ операций умножения и $(N/2)(N/2 - 1) \cdot 2 = N(N - 2)/2$ операций сложения. Так как расчет ДПФ $X(jk\Omega)$ исходной N -точечной последовательности из ДПФ $X_1(jk\Omega)$ и $X_2(jk\Omega)$ требует небольшого числа операций, то однократное применение указанной процедуры позволяет сократить объем вычислений практически вдвое.

Для дальнейшего уменьшения числа операций описанную процедуру применяют многократно. Если $N = 2^r$, где r — целое число, то минимальный размер для вспомогательных последовательностей при расчете ДПФ будет двухточечным. При этом число ступеней вычислений составляет $r = \log_2 N$, общее количество операций умножения и сложения, требующееся для расчета искомого ДПФ, составляет примерно $2N \log_2 N$, а выигрыш в количестве требуемых операций по отношению к выражению (1.14) — $\frac{N}{\log_2 N}$ раз.

Особо следует учесть, что БПФ представляет собой не новый вид преобразования Фурье, а целый ряд эффективных алгоритмов, предназначенных для быстрого вычисления ДПФ. Наиболее распространенными алгоритмами БПФ являются алгоритмы с прореживанием по времени и по частоте.

1.13.1. Быстрое преобразование Фурье с прореживанием по времени

Ниже представлен алгоритм БПФ с прореживанием по времени, причем для простоты записи использовано обозначение ДПФ в виде $X(k)$.

Пусть исходная N -точечная последовательность $x(n)$ имеет четное число отсчетов, причем $N = 2^r$, где r — целое число. Разделим исходную последовательность на две вспомогательные $N/2$ -точечные последовательности

$$x_1(n) = x(2n) \text{ и } x_2(n) = x(2n + 1) \quad \text{при } n = 0, 1, \dots, \frac{N}{2} - 1.$$

Очевидно, что $x_1(n)$ содержит в себе четные члены исходной последовательности, а $x_2(n)$ — нечетные. На основании выражения (1.14) можно записать следующее:

$$X(k) = \underbrace{\sum_{n=0}^{N/2-1} x_1(n) W_N^{kn}}_{n - \text{четные}} + \underbrace{\sum_{n=0}^{N/2-1} x_2(n) W_N^{kn}}_{n - \text{нечетные}}.$$

Тогда с учетом вспомогательных последовательностей $x_1(n)$ и $x_2(n)$ N -точечное ДПФ для сигнала $x(n)$ можно представить в виде

$$\begin{aligned} X(k) &= \sum_{n=0}^{N/2-1} x(2n)W_N^{2nk} + \sum_{n=0}^{N/2-1} x(2n+1)W_N^{(2n+1)k} = \\ &= \sum_{n=0}^{N/2-1} x(2n)(W_N^2)^{nk} + W_N^k \sum_{n=0}^{N/2-1} x(2n+1)(W_N^2)^{nk}. \end{aligned}$$

Так как $W_N^2 = (e^{-j(2\pi/N)})^2 = e^{-j(2\pi/(N/2))} = W_{N/2}$, то

$$X(k) = \sum_{n=0}^{N/2-1} x(2n)W_{N/2}^{nk} + W_N^k \sum_{n=0}^{N/2-1} x(2n+1)W_{N/2}^{nk}. \quad (1.15)$$

Суммы в правой части полученного выражения представляют собой $N/2$ -точечные ДПФ вспомогательных последовательностей:

$$\begin{aligned} X_1(k) &= \sum_{n=0}^{N/2-1} x(2n)W_{N/2}^{nk} \quad \text{при } k = 0, 1, 2, \dots, \frac{N}{2} - 1; \\ X_2(k) &= \sum_{n=0}^{N/2-1} x(2n+1)W_{N/2}^{nk} \quad \text{при } k = 0, 1, 2, \dots, \frac{N}{2} - 1. \end{aligned}$$

С учетом этого выражение (1.15) можно представить в виде

$$X(k) = X_1(k) + W_N^k X_2(k). \quad (1.16)$$

Поскольку компоненты $X(k)$ должны быть определены для $k = 0, 1, 2, \dots, N-1$, необходимо доопределить выражение (1.16) для $k = N/2, N/2+1, N/2+2, \dots, N-1$. Это легко сделать, используя свойство периодичности ДПФ. Так как $N/2$ -точечное ДПФ периодически с периодом в $N/2$ точек, то

$$X_1(k) = X_1\left(k + \frac{N}{2}\right); \quad X_2(k) = X_2\left(k + \frac{N}{2}\right).$$

Тогда из формулы (1.16) следует, что

$$X\left(k + \frac{N}{2}\right) = X_1\left(k + \frac{N}{2}\right) + W_N^{k+N/2} X_2\left(k + \frac{N}{2}\right). \quad (1.17)$$

Поскольку

$$\begin{aligned} W_N^{k+N/2} &= \exp\left(-j2\pi \frac{k+N/2}{N}\right) = \exp(-j2\pi k/N) \exp(-j\pi) = \\ &= -\exp(-j2\pi k/N) = -W_N^k, \end{aligned}$$

то уравнение (1.17) можно представить в виде

$$X\left(k + \frac{N}{2}\right) = X_1(k) - W_N^k X_2(k).$$

Следовательно, N -точечное ДПФ $X(k)$ полностью определяется следующим выражением:

$$X(k) = \begin{cases} X_1(k) + W_N^k X_2(k) & \text{при } k = 0, 1, \dots, \frac{N}{2} - 1, \\ X_1(k) - W_N^k X_2(k) & \text{при } k = \frac{N}{2}, \frac{N}{2} + 1, \dots, N - 1. \end{cases} \quad (1.18)$$

Выражение (1.18) определяет базовую операцию БПФ с прореживанием по времени, называемую «бабочкой» и представляемую графически в виде ориентированного графа (рис. 1.29). В его центральном узле осуществляются операции сложения и вычитания, причем верхний выход соответствует сумме входов, а нижний — разности. Стрелка графа обозначает умножение на сомножитель W_N^k .

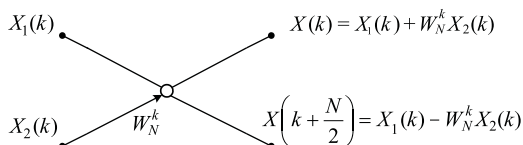


Рис. 1.29. Графическое представление базовой операции «бабочка» алгоритма БПФ с прореживанием по времени

Существенно, что $N/2$ -точечные ДПФ $X_1(k)$ и $X_2(k)$, входящие в уравнение (1.18), также могут быть рассчитаны алгоритмом БПФ через вспомогательные $N/4$ -точечные ДПФ и т. д. Максимальная эффективность рассматриваемого алгоритма обеспечивается при делении анализируемого сигнала $x(n)$ на 2-точечные вспомогательные. Графическое представление процедуры расчета спектра 8-точечной дискретной последовательности на основании алгоритма БПФ с прореживанием по времени показано на рис. 1.30. В этом случае дробление исходной последовательности $x(n)$ выполняется в два этапа. На первом этапе формируются вспомогательные последовательности: $x_1(n) = \{x(0), x(2), x(4), x(6)\}$, состоящая из четных членов исходного сигнала, и $x_2(n) = \{x(1), x(3), x(5), x(7)\}$, содержащая нечетные члены.

В результате второго этапа дробления формируются 2-точечные вспомогательные дискретные сигналы $x_{11}(n) = \{x(0), x(4)\}$, $x_{12}(n) = \{x(2), x(6)\}$, $x_{21}(n) = \{x(1), x(5)\}$ и $x_{22}(n) = \{x(3), x(7)\}$. Указанный способ формирования вспомогательных сигналов в алгоритме

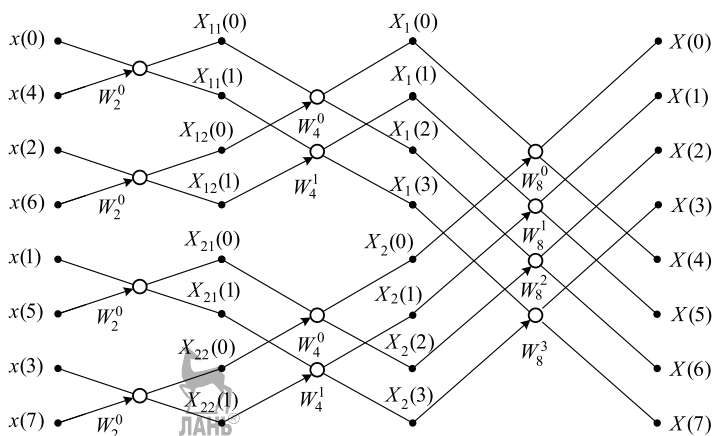


Рис. 1.30. Расчет спектра 8-точечной дискретной последовательности алгоритмом быстрого преобразования Фурье с прореживанием по времени

БПФ с прореживанием по времени обеспечивает правильную очередность формирования спектральных отсчетов $X(k)$ ДПФ исходного сигнала и часто представляется процедурой двоичной инверсии номеров, которая определяется следующим образом. Порядковый номер n отсчета исходного сигнала записывается в форме двоичной r -разрядной комбинации. Затем порядок следования разрядов инвертируется (старший разряд становится младшим и т. д.). Десятичное число, соответствующее новой двоичной комбинации, и определяет номер отсчета исходной комбинации после перестановки («прореживания»). В табл. 1.3 представлена процедура двоичной инверсии номеров для случая 8-точечного исходного сигнала $x(n)$.

Из рис. 1.30 видно, что для обеспечения «нужной» очередности формирования k -х частотных составляющих в алгоритме БПФ с

Таблица 1.3

Процедура двоичной инверсии номеров отсчетов входного сигнала

Отсчет исходного сигнала $x(n)$	Кодовая комбинация номера отсчета	Инвертированная кодovая комбинация номера отсчета	Десятичный номер отсчета после двоичной инверсии
$x(0)$	000	000	0
$x(1)$	001	100	4
$x(2)$	010	010	2
$x(3)$	011	110	6
$x(4)$	100	001	1
$x(5)$	101	101	5
$x(6)$	110	011	3
$x(7)$	111	111	7

прореживанием по времени изменяется очередность использования n -х отсчетов входной последовательности.

Пример 1.15. С использованием алгоритма быстрого преобразования Фурье с прореживанием по времени требуется произвести спектральное оценивание периодического дискретного сигнала $x(nT) = \{1, 1, 1, 0, 0, 0\}$, имеющего период дискретизации $T = 1$ мс.

Следует отметить, что решение этого примера традиционным алгоритмом ДПФ (1.14) представлено выше (пример 1.14), что позволяет произвести сравнение полученных результатов.

Период исследуемого сигнала содержит количество отсчетов, не кратное 2 ($N = 6$), что не позволяет добиться максимальной эффективности БПФ. В этом случае решение задачи возможно следующим образом:

- первый этап — разделение сигнала $x(nT)$ на вспомогательные сигналы $x_1(nT_1) = \{x(0), x(2), x(4)\}$ и $x_2(nT_2) = \{x(1), x(3), x(5)\}$ (прореживание по времени);
- второй этап — расчет спектров $X_1(jk\Omega_1)$ и $X_2(jk\Omega_2)$ вспомогательных сигналов $x_1(nT_1)$ и $x_2(nT_2)$ процедурой ДПФ;
- третий этап — расчет искомого спектра $X(jk\Omega)$ исходного сигнала $x(nT)$ процедурой БПФ в соответствии с формулой (1.18).

Графическая иллюстрация вычислительных процедур представлена на рис. 1.31.

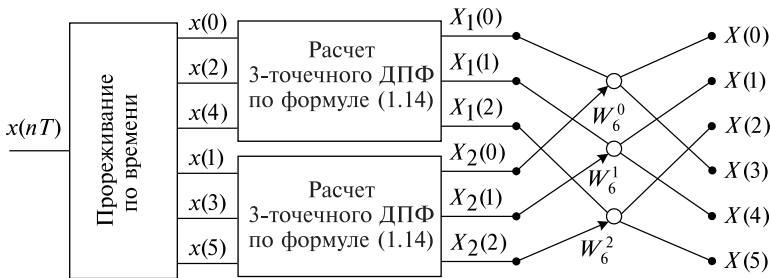


Рис. 1.31. Структура вычислительных процедур к примеру 1.15

Первый этап. Процедура прореживания по времени обуславливает формирование периодических вспомогательных сигналов $x_1(nT_1) = \{1, 1, 0\}$ и $x_2(nT_2) = \{1, 0, 0\}$ со следующими параметрами: $N_1 = N_2 = 3$, $T_1 = T_2 = 2$ мс. Графики сигналов $x(nT)$, $x_1(nT_1)$ и $x_2(nT_2)$ показаны на рис. 1.32.

Второй этап. Спектр $X_1(jk\Omega_1)$ вспомогательного сигнала $x_1(nT_1)$ определяется дискретным преобразованием Фурье в соответствии с выражением (1.14):

$$X_1(k) = \sum_{n=0}^{N_1-1} x_1(n) e^{-j(2\pi/N_1)kn} =$$

$$= x_1(0) e^{-j(2\pi/N_1)k \cdot 0} + x_1(1) e^{-j(2\pi/N_1)k \cdot 1} + x_1(2) e^{-j(2\pi/N_1)k \cdot 2} = 1 + e^{-j(2\pi/3)k}.$$

Следовательно, спектральные составляющие имеют вид

$$X_1(0) = 1 + e^{-j(2\pi/3) \cdot 0} = 1 + 1 = 2 = 2e^{j0};$$

$$X_1(1) = 1 + e^{-j(2\pi/3) \cdot 1} = 1 + \cos \frac{2\pi}{3} - j \sin \frac{2\pi}{3} = 1 - \frac{1}{2} - j \frac{\sqrt{3}}{2} = \frac{1}{2} - j \frac{\sqrt{3}}{2} = e^{-j\pi/3};$$

$$X_1(2) = 1 + e^{-j(2\pi/3) \cdot 2} = 1 + \cos \frac{4\pi}{3} - j \sin \frac{4\pi}{3} = 1 - \frac{1}{2} + j \frac{\sqrt{3}}{2} = \frac{1}{2} + j \frac{\sqrt{3}}{2} = e^{j\pi/3}.$$

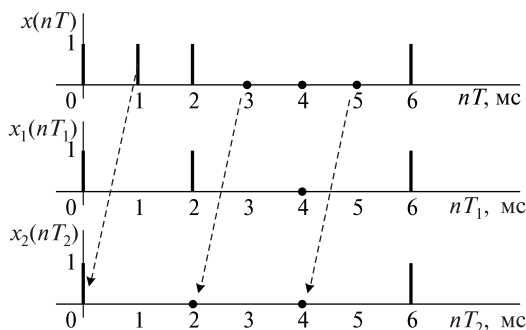


Рис. 1.32. Графики исходного и вспомогательных сигналов

Спектр $X_2(jk\Omega_2)$ вспомогательного сигнала $x_2(nT_2)$ определяется дискретным преобразованием Фурье:

$$X_2(k) = \sum_{n=0}^{N_2-1} x_2(n) e^{-j(2\pi/N_2)kn} =$$

$$= x_2(0) e^{-j(2\pi/N_2)k \cdot 0} + x_2(1) e^{-j(2\pi/N_2)k \cdot 1} + x_2(2) e^{-j(2\pi/N_2)k \cdot 2} = 1.$$

Следовательно, спектральные составляющие имеют вид

$$X_2(0) = X_2(1) = X_2(2) = 1 = e^{j0}.$$

Основные параметры спектральных составляющих сигналов $x_1(nT_1)$ и $x_2(nT_2)$, полученных по результатам расчетов, представлены в табл. 1.4. Существенно, что основные частоты ДПФ Ω_1 и Ω_2 (бины ДПФ) для спектров $X_1(jk\Omega_1)$ и $X_2(jk\Omega_2)$ одинаковы и равны основной частоте ДПФ искомого спектра $X(jk\Omega)$:

$$\Omega_1 = \Omega_2 = \frac{2\pi}{N_1 T_1} = \frac{2\pi}{N_2 T_2} = \frac{2\pi}{3 \cdot 2 \cdot 10^{-3}} = \frac{10^3 \pi}{3} \text{ рад/с} = \Omega.$$

Таблица 1.4

Основные параметры спектральных составляющих вспомогательных сигналов

Спектральные составляющие	Амплитуда $A(k\Omega) = X(jk\Omega) $	Начальная фаза $\varphi(k\Omega) = \arg X(jk\Omega)$, рад	Угловая частота $\omega_k = k\Omega$, рад/с	Частота $f_k = \frac{\omega_k}{2\pi}$, Гц	Период $T_k = \frac{1}{f_k}$, мс
$X_1(0)$	2	0	0	0	0
$X_1(1)$	1	$-\frac{\pi}{3}$	$\frac{10^3 \pi}{3}$	$\frac{10^3}{6}$	6
$X_1(2)$	1	$\frac{\pi}{3}$	$\frac{2 \cdot 10^3 \pi}{3}$	$\frac{10^3}{3}$	3
$X_2(0)$	1	0	0	0	0
$X_2(1)$	1	0	$\frac{10^3 \pi}{3}$	$\frac{10^3}{6}$	6
$X_2(2)$	1	0	$\frac{2 \cdot 10^3 \pi}{3}$	$\frac{10^3}{3}$	3

Третий этап. Искомый спектр $X(jk\Omega)$ анализируемого сигнала $x(nT)$ определяется на основании БПФ с прореживанием по времени:

$$X(k) = \begin{cases} X_1(k) + W_N^k X_2(k) & \text{при } k = 0, 1, \dots, \frac{N}{2} - 1, \\ X_1(k) - W_N^k X_2(k) & \text{при } k = \frac{N}{2}, \frac{N}{2} + 1, \dots, N - 1. \end{cases}$$

Ниже представлен расчет спектральных составляющих $X(k)$ в последовательности, соответствующей очередности расположения «бабочек» на рис. 1.31.

$$X(0) = X_1(0) + W_6^0 X_2(0) = 2e^{j0} + e^{-j(2\pi/6)0} e^{j0} = 2 + 1 = 3 = 3e^{j0};$$

$$X(3) = X_1(0) - W_6^0 X_2(0) = 2e^{j0} - e^{-j(2\pi/6)0} e^{j0} = 2 - 1 = 1 = 1e^{j0};$$

$$X(1) = X_1(1) + W_6^1 X_2(1) = e^{-j\pi/3} + e^{-j(2\pi/6)1} e^{j0} = e^{-j\pi/3} + e^{-j\pi/3} = 2e^{-j\pi/3};$$

$$X(4) = X_1(1) - W_6^1 X_2(1) = e^{-j\pi/3} - e^{-j(2\pi/6)1} e^{j0} = e^{-j\pi/3} - e^{-j\pi/3} = 0;$$

$$X(2) = X_1(2) + W_6^2 X_2(2) = e^{j\pi/3} + e^{-j(2\pi/6)2} e^{j0} = e^{j\pi/3} + e^{-j(2\pi/3)} =$$

$$= \cos \frac{\pi}{3} + j \sin \frac{\pi}{3} + \cos \frac{2\pi}{3} - j \sin \frac{2\pi}{3} = \frac{1}{2} + j \frac{\sqrt{3}}{2} + \left(-\frac{1}{2}\right) - j \frac{\sqrt{3}}{2} = 0;$$

$$X(5) = X_1(2) - W_6^2 X_2(2) = e^{j\pi/3} - e^{-j(2\pi/6)2} e^{j0} = e^{j\pi/3} - e^{-j(2\pi/3)} =$$

$$= \cos \frac{\pi}{3} + j \sin \frac{\pi}{3} - \cos \frac{2\pi}{3} + j \sin \frac{2\pi}{3} = \frac{1}{2} + j \frac{\sqrt{3}}{2} - \left(-\frac{1}{2}\right) + j \frac{\sqrt{3}}{2} = 1 + j\sqrt{3} = 2e^{j\pi/3}.$$

Полученные результаты полностью соответствуют спектральным составляющим сигнала $x(nT)$, рассчитанным по традиционному вычислительному алгоритму ДПФ.

1.13.2. Быстрое преобразование Фурье с прореживанием по частоте

В рассматриваемом алгоритме, в отличие от алгоритма БПФ с прореживанием по времени, очередность следования отсчетов анализируемого сигнала $x(n)$ не изменяется, а формирование спектральных отсчетов $X(k)$ происходит в последовательности, требующей их перестановки (прореживание по частоте).

Пусть исходная N -точечная последовательность $x(n)$ имеет четное число отсчетов, причем $N = 2^r$, где r — целое число. В основе алгоритма БПФ с прореживанием по частоте лежит разделение исходной последовательности на две вспомогательные $N/2$ -точечные последовательности $x_1(n)$ и $x_2(n)$ следующим образом:

$$x_1(n) = x(n) \quad \text{при } n = 0, 1, \dots, N/2 - 1;$$

$$x_2(n) = x(n) \quad \text{при } n = \frac{N}{2}, \frac{N}{2} + 1, \frac{N}{2} + 2, \dots, N - 1.$$

Очевидно, что $x_1(n)$ представляет собой первую половину исходного сигнала, а $x_2(n)$ — его вторую половину.

На основании формулы (1.14) ДПФ анализируемого сигнала может быть определено выражением

$$X(k) = \sum_{n=0}^{N/2-1} x(n) e^{-j(2\pi/N)nk} + \sum_{n=\frac{N}{2}}^{N-1} x(n) e^{-j(2\pi/N)nk}. \quad (1.19)$$

С учетом вспомогательных сигналов и соотношения $x_2(n) = x(n + N/2)$ при $n = 0, 1, \dots, N/2 - 1$ выражение (1.19) преобразуется к виду

$$\begin{aligned} X(k) &= \sum_{n=0}^{N/2-1} x(n) e^{-j(2\pi/N)nk} + \sum_{n=0}^{N/2-1} x\left(n + \frac{N}{2}\right) e^{-j(2\pi/N)(n+N/2)k} = \\ &= \sum_{n=0}^{N/2-1} x(n) W_N^{nk} + \sum_{n=0}^{N/2-1} x\left(n + \frac{N}{2}\right) W_N^{(n+N/2)k}. \end{aligned} \quad (1.20)$$

Поворачивающий множитель $W_N^{(n+N/2)k}$ может быть представлен произведением $W_N^{(n+N/2)k} = W_N^{nk} W_N^{(N/2)k}$, второй сомножитель которого определяется выражением

$$W_N^{(N/2)k} = e^{-j(2\pi/N)(N/2)k} = e^{-j\pi k} = (-1)^k$$

и принимает значение «1» для всех четных спектральных составляющих спектра $X(k)$ и значение «-1» для его нечетных спектральных составляющих. Следовательно, формула (1.20) может быть преобразована к виду

$$X(k) = \sum_{n=0}^{N/2-1} [x_1(n) + x_2(n)(-1)^k] W_N^{nk}. \quad (1.21)$$

Подстановка в выражение (1.21) вместо k переменных $2k$ и $2k+1$ позволяет сформировать упрощенные уравнения для расчета четных $X(2k)$ и нечетных $X(2k+1)$ спектральных составляющих:

$$\begin{aligned} X(2k) &= \sum_{n=0}^{N/2-1} [x_1(n) + x_2(n)] W_N^{n2k} = \sum_{n=0}^{N/2-1} [x_1(n) + x_2(n)] W_{N/2}^{nk}; \\ X(2k+1) &= \sum_{n=0}^{N/2-1} [x_1(n) - x_2(n)] W_N^{n(2k+1)} = \\ &= \sum_{n=0}^{N/2-1} \{[x_1(n) - x_2(n)] W_N^n\} W_N^{n2k} = \sum_{n=0}^{N/2-1} \{[x_1(n) - x_2(n)] W_N^n\} W_{N/2}^{nk}. \end{aligned}$$

Следовательно, N -точечное ДПФ $X(k)$ при расчете алгоритмом БПФ с прореживанием по частоте формируется парой выражений:

$$\begin{aligned} X(2k) &= \sum_{n=0}^{N/2-1} [x_1(n) + x_2(n)] W_{N/2}^{nk} \quad \text{при } k = 0, 1, \dots, \frac{N}{2} - 1; \\ X(2k+1) &= \sum_{n=0}^{N/2-1} \{[x_1(n) - x_2(n)] W_N^n\} W_{N/2}^{nk} \quad \text{при } k = 0, 1, \dots, \frac{N}{2} - 1. \end{aligned} \quad (1.22)$$

представляющих собой $N/2$ -точечные ДПФ суммы и разности сигналов $x_1(n)$ и $x_2(n)$. Разность перед вычислением ДПФ умножается на множитель W_N^n .

Выражение (1.22) можно представить как

$$X(2k) = X_1(k) + X_2(k) \quad \text{при } k = 0, 1, 2, \dots, \frac{N}{2} - 1;$$

$$X(2k+1) = [X_1(k) - X_2(k)]W_N^n \quad \text{при } k = 0, 1, 2, \dots, \frac{N}{2} - 1,$$

где $X_1(k) = \sum_{n=0}^{N/2-1} x_1(n)W_{N/2}^{nk}$ и $X_2(k) = \sum_{n=0}^{N/2-1} x_2(n)W_{N/2}^{nk}$.

В связи с этим «бабочка» быстрого преобразования Фурье с прореживанием по частоте имеет структуру, показанную на рис. 1.33. В центральном узле графа осуществляются операции сложения и вычитания, причем верхний выход соответствует сумме входов, а нижний — разности. Стрелка графа означает умножение на сомножитель W_N^n .

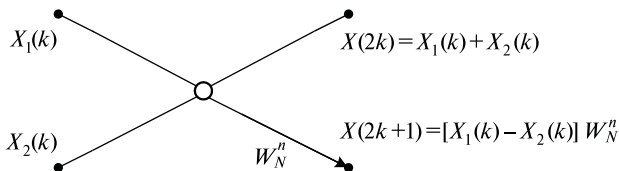


Рис. 1.33. Графическое представление базовой операции «бабочка» алгоритма БПФ с прореживанием по частоте

Очевидно, что при расчете спектральных составляющих сигналов, содержащих $N = 2^r$ отсчетов, операция «бабочка» может быть выполнена последовательно r раз, что обеспечивает максимальную эффективность алгоритма по количеству вычислительных операций. В отличие от алгоритма БПФ с прореживанием по времени, в рассматриваемом алгоритме на каждом этапе расчетов получаемое ДПФ определяется через ДПФ вдвое большей размерности.

Графическое представление процедуры расчета спектра 8-точечной дискретной последовательности на основании алгоритма БПФ с прореживанием по частоте показано на рис. 1.34.

В этом случае первоначальное дробление исходной последовательности $x(n)$ заключается в формировании вспомогательных последовательностей: $x_1(n) = \{x(0), x(1), x(2), x(3)\}$, представляющей первую половину исходного сигнала, и $x_2(n) = \{x(4), x(5), x(6), x(7)\}$, представляющей вторую его половину.

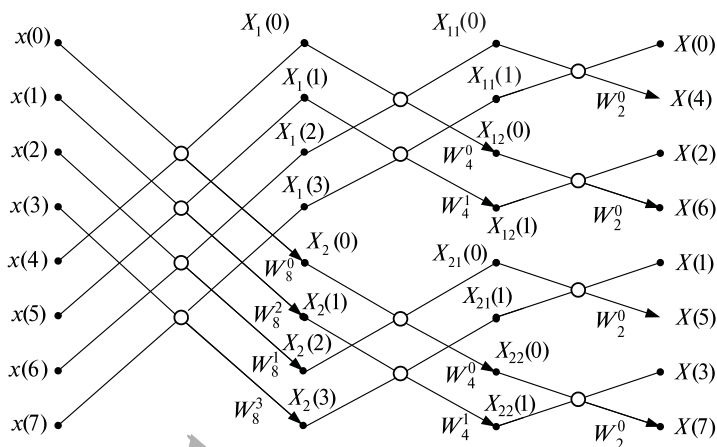


Рис. 1.34. Расчет спектра 8-точечной дискретной последовательности алгоритмом быстрого преобразования Фурье с прореживанием по частоте

Характерно, что на этапе обработки двухточечных последовательностей умножение на сомножитель W_N^n практически не нужно, так как $W_2^0 = 1$. На этом этапе рассчитываемые спектральные отсчеты представляют собой сумму и разность исходных отсчетов. Например, по отношению к рис. 1.34 справедливы равенства $X(0) = X_{11}(0) + X_{11}(1)$ и $X(4) = X_{11}(0) - X_{11}(1)$.

Способ формирования вспомогательных последовательностей, используемый в алгоритме БПФ с прореживанием по частоте, обеспечивает очередность формирования спектральных отсчетов $X(k)$ ДПФ исходного сигнала, требующую выполнения процедуры двоичной инверсии номеров.

Порядковый номер спектральной составляющей исходного сигнала, записанный в форме двоичной r -разрядной комбинации, должен быть инвертирован (старший разряд становится младшим и т. д.).

Десятичное число, соответствующее новой полученной двоичной комбинации, определяет номер k спектральной составляющей после перестановки (прореживание по частоте) (табл. 1.5).

Пример 1.16. С использованием БПФ с прореживанием по частоте произвести спектральное оценивание периодического дискретного сигнала $x(nT) = \{1, 1, 1, 0, 0, 0\}$ с периодом дискретизации $T = 1$ мс.

Ранее спектральная оценка указанного сигнала была выполнена при помощи традиционного алгоритма дискретного преобразования Фурье и алгоритма БПФ с прореживанием по времени (примеры 1.14 и 1.15), что позволяет сравнивать полученные результаты.

Количество отсчетов, содержащихся на периоде анализируемого сигнала, не равно целой степени числа «2» ($N = 6$), что не создает условий для достижения мак-

Таблица 1.5

Процедура двоичной инверсии порядковых номеров спектральных составляющих анализируемого сигнала

Спектральная составляющая $X(k)$ в порядке формирования	Кодовая комбинация порядкового номера спектральной составляющей	Инвертированная кодовая комбинация порядкового номера спектральной составляющей	Десятичный номер спектральной составляющей после двоичной инверсии порядкового номера
$X(0)$	000	000	0
$X(4)$	001	100	4
$X(2)$	010	010	2
$X(6)$	011	110	6
$X(1)$	100	001	1
$X(5)$	101	101	5
$X(3)$	110	011	3
$X(7)$	111	111	7

симальной эффективности процедуры БПФ. Следовательно, возможно однократное дробление сигнала $x(nT)$ на вспомогательные последовательности: $x_1(n) = \{1, 1, 1\}$, представляющую первую половину исходного сигнала, и $x_2(n) = \{0, 0, 0\}$, представляющую вторую его половину.

Спектральные составляющие с четными номерами рассчитываются, используя первое уравнение системы:

$$X(2k) = \sum_{n=0}^{N/2-1} [x_1(n) + x_2(n)]W_{N/2}^{nk} \quad \text{при } k = 0, 1, \dots, \frac{N}{2} - 1.$$

Соответствующие дискретные гармоники имеют вид

$$\begin{aligned} X(0) &= [x_1(0) + x_2(0)]W_3^{0 \cdot 0} + [x_1(1) + x_2(1)]W_3^{1 \cdot 0} + [x_1(2) + x_2(2)]W_3^{2 \cdot 0} = \\ &= (1 + 0) \cdot 1 + (1 + 0) \cdot 1 + (1 + 0) \cdot 1 = 3; \\ X(2) &= [x_1(0) + x_2(0)]W_3^{0 \cdot 1} + [x_1(1) + x_2(1)]W_3^{1 \cdot 1} + [x_1(2) + x_2(2)]W_3^{2 \cdot 1} = \\ &= 1 + e^{-j(2\pi/3)} + e^{-j4\pi/3} = 0; \\ X(4) &= [x_1(0) + x_2(0)]W_3^{0 \cdot 2} + [x_1(1) + x_2(1)]W_3^{1 \cdot 2} + [x_1(2) + x_2(2)]W_3^{2 \cdot 2} = \\ &= 1 + e^{-j4\pi/3} + e^{-j8\pi/3} = 0. \end{aligned}$$

Спектральные составляющие с нечетными номерами рассчитываются, используя второе уравнение системы:

$$X(2k+1) = \sum_{n=0}^{N/2-1} \{[x_1(n) - x_2(n)]W_N^n\}W_{N/2}^{nk} \quad \text{при } k = 0, 1, \dots, \frac{N}{2} - 1.$$

Соответствующие дискретные гармоники имеют вид

$$\begin{aligned} X(1) &= [x_1(0) - x_2(0)]W_6^0 W_3^{0 \cdot 0} + [x_1(1) - x_2(1)]W_6^1 W_3^{1 \cdot 0} + [x_1(2) - x_2(2)]W_6^2 W_3^{2 \cdot 0} = \\ &= 1 + e^{-j(2\pi/6)} + e^{-j2\pi/3} = 2e^{-j\pi/3}; \\ X(3) &= [x_1(0) - x_2(0)]W_6^0 W_3^{0 \cdot 1} + [x_1(1) - x_2(1)]W_6^1 W_3^{1 \cdot 1} + [x_1(2) - x_2(2)]W_6^2 W_3^{2 \cdot 1} = \\ &= 1 + e^{-j(2\pi/6)} e^{-j(2\pi/3)} + e^{-j4\pi/6} e^{-j4\pi/3} = 1 + e^{-j\pi} + e^{-j2\pi} = 1; \\ X(5) &= [x_1(0) - x_2(0)]W_6^0 W_3^{0 \cdot 2} + [x_1(1) - x_2(1)]W_6^1 W_3^{1 \cdot 2} + [x_1(2) - x_2(2)]W_6^2 W_3^{2 \cdot 2} = \\ &= 1 + e^{-j(2\pi/6)} e^{-j4\pi/3} + e^{-j4\pi/6} e^{-j8\pi/3} = 1 + e^{-j5\pi/3} + e^{-j10\pi/3} = 2e^{j\pi/3}. \end{aligned}$$

Полученные результаты соответствуют спектральным составляющим сигнала $x(nT)$, рассчитанным алгоритмами дискретного преобразования Фурье и БПФ с проживанием по времени.

1.14. Дискретное косинусное преобразование

Дискретное косинусное преобразование (ДКП) тесно связано с ДПФ (особенно с его действительной частью) и предназначено для спектрального разложения действительных периодических (или конечных) четных сигналов. Очевидно, что спектр таких сигналов состоит лишь из косинусоидальных компонентов, что обуславливает название преобразования.

В состав базисной системы ДКП (ядра ДКП) входят ортогональные действительные функции

$$\varphi_k = \cos\left(\frac{2n+1}{2N}k\pi\right), \quad k, n = 0, 1, 2, \dots, N-1.$$

Существенно, что базисные функции N -точечного ДКП представляют собой действительную часть произведения ядра $2N$ -точечного ДПФ на множитель $\exp(-j(\pi/2N)k)$:

$$\begin{aligned} \varphi_k &= \operatorname{Re}\left\{W_{2N}^{kn} \exp\left(-j\frac{\pi}{2N}k\right)\right\} = \operatorname{Re}\left\{\exp\left(-j\frac{2\pi}{2N}kn\right) \times \right. \\ &\times \left.\exp\left(-j\frac{\pi}{2N}k\right)\right\} = \operatorname{Re}\left\{\exp\left(-j\frac{2n+1}{2N}k\pi\right)\right\} = \cos\left(\frac{2n+1}{2N}k\pi\right). \end{aligned}$$

В общем виде прямое N -точечное ДКП представляется как

$$X_c(k) = \sqrt{\frac{2}{N}} C(k) \sum_{n=0}^{N-1} x(n) \cos\left[\frac{(2n+1)k\pi}{2N}\right]; \quad (1.23)$$

$$C(k) = \begin{cases} 1/\sqrt{2}, & k = 0; \\ 1, & k = 1, 2, \dots, N-1. \end{cases} \quad (1.24)$$

Обратное дискретное косинусное преобразование определяется следующей формулой:

$$x(n) = \sqrt{\frac{2}{N}} \sum_{k=0}^{N-1} C(k) X_c(k) \cos\left[\frac{(2n+1)k\pi}{2N}\right].$$

Следует иметь в виду, что реальные сигналы, являясь, как правило, действительными, в общем случае не соответствуют требованиям периодичности и четности. Это обстоятельство обуславливает необходимость корректного использования процедуры ДКП. Способы представления непериодического и нечетного сигнала, обеспечивающие правомерность применения ДПФ (БПФ) и ДКП, представлены на рис. 1.35.

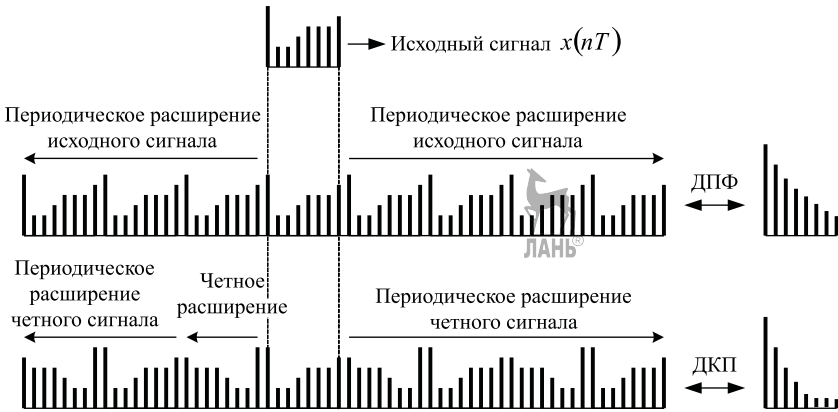


Рис. 1.35. Представление сигналов для корректного применения ДПФ и ДКП

В практике применения ДКП выполнение операций четного и периодического расширений не обязательно, достаточно правильного применения результатов преобразования.

Преимуществом ДКП по сравнению с ДПФ является большая скорость затухания спектральных составляющих (качественно это показано в правой части рис. 1.35), что означает сосредоточение значительной доли энергии в меньшем числе спектральных коэффициентов и делает возможным достаточно точное описание исходного сигнала $x(nT)$ меньшим числом трансформант $X_c(k)$.

Следует иметь в виду, что для практического применения разработаны алгоритмы ДКП, основанные на БПФ, а также алгоритмы быстрого ДКП (БДКП), скорость вычисления которого в несколько раз выше скорости вычисления ДКП.

Основной областью применения дискретного косинусного преобразования в современных системах инфокоммуникаций является трансформное кодирование звуковых сигналов и сигналов изображений.

В задачах кодирования изображений используется, как правило, двумерное ДКП блока отсчетов размером $N \times N$ (например, в стандартах JPEG и MPEG $N = 8$). В общем случае при яркости $x(n, m)$ пикселей с координатами n и m соответственно коэффициенты $X_c(k, l)$ двумерного ДКП с координатами k и l определяются выражением

$$X_c(k, l) = \frac{2}{N} C(k) C(l) \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} x(n, m) \cos \left[\frac{(2n+1)l\pi}{2N} \right] \cos \left[\frac{(2m+1)k\pi}{2N} \right],$$

в котором множитель $C(l)$ вычисляется аналогично выражению (1.24). Обратное двумерное ДКП определяется по формуле

$$x(n, m) = \frac{2}{N} \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} C(k)C(l)X_c(k, l) \cos \left[\frac{(2n+1)l\pi}{2N} \right] \cos \left[\frac{(2m+1)k\pi}{2N} \right].$$

Двумерное ДКП является сепарабельным (разделимым) преобразованием и может реализовываться последовательным выполнением одномерного ДКП (1.23) в каждом из измерений.

Различные модификации ДКП применяются в стандартах кодирования звуковых сигналов:

- для кодирования речевого сигнала с диапазоном частот 50...7000 Гц в соответствии со стандартом G.722.1 Международного союза электросвязи используется алгоритм быстрого ДКП (часто называемый «ДКП формы IV»);
- модифицированный вариант ДКП (МДКП, MDCT — Modified Discrete Cosine Transform) используется при кодировании широкополосных звуковых сигналов в следующих стандартах:
 1. MPEG-1 ISO/IEC 11172-3 (уровень 3) — для кодирования звукового сигнала с полосой частот 40-15000 Гц в системах Интернет-вещания.
 2. MPEG-2 ISO/IEC 13818-3, расширяющем стандарт MPEG-1 и используемом в системах многоканального стереофонического вещания.
 3. MPEG-2 ISO/IEC 13818-7 — для кодирования звукового сигнала с полосой частот до 20 кГц на основе алгоритма AAC (Advanced Audio Coding) в системах многоканального стереофонического вещания.
 4. MPEG-4, ориентированном на мультимедиаприложения.
 5. Dolby AC-3, рекомендованном американским национальным комитетом ATSC (Advanced Television System Committee) для систем телевидения высокой четкости, кинематографических систем, систем спутникового вещания, систем записи звуковых сигналов на магнитные, оптические и другие носители информации.

1.15. Дискретное преобразование Лапласа.

Z-преобразование дискретных сигналов и его свойства

В теории обработки непрерывных сигналов $x(t)$ большую популярность имеют прямое и обратное преобразования Лапласа, опре-

деляемые парой соотношений с использованием комплексной переменной p :

$$X(p) = \int_0^{\infty} x(t)e^{-pt} dt;$$

$$x(t) = \frac{1}{2\pi j} \int_{\sigma-j\omega}^{\sigma+j\omega} X(p)e^{pt} dp.$$

Очевидно, что преобразование Лапласа является обобщением непрерывного преобразования Фурье, при этом исследование исходного сигнала производится не только на мнимой оси $j\omega$, а во всей комплексной плоскости $p = \sigma + j\omega$, что позволяет получить больше необходимой информации. Основным вычислительным удобством преобразования Лапласа является равносильная замена операции дифференцирования функции $x(t)$ по времени в пространстве оригиналов операцией умножения функции $X(p)$ на переменную p в пространстве изображений, что позволяет свести дифференциальные вычисления к алгебраическим.

Отмеченные преимущества определяют достаточно частое применение в теории ЦОС дискретного преобразования Лапласа (ДПЛ), являющегося распространением преобразования Лапласа на случай дискретных сигналов. Прямое ДПЛ представляет собой функциональное преобразование решетчатой функции $x(nT)$ вида

$$X(p) = \sum_{n=0}^{\infty} x(nT)e^{-pnT} \quad (1.25)$$

и позволяет производить над разностными уравнениями, используемыми в ЦОС, такие же алгебраические действия, какие допускает обычное преобразование Лапласа над интегрально-дифференциальными уравнениями в теории непрерывных сигналов. Особенностью ДПЛ является то, что p -изображения состоят из функций e^{-pT} . Но поскольку функция e^{-pT} является периодической вдоль мнимой оси p -плоскости с периодом $2\pi/T$, т. е. $e^{-pT} = e^{-(pT \pm j(2\pi k/T)T)}$, $k = 0, 1, 2, \dots$, то изображения $X(p)$ полностью определяются в любой полосе p -плоскости, параллельной действительной оси и имеющей ширину $2\pi/T$. Обычно эта полоса выбирается симметрично действительной оси ($-\pi/T < \omega \leq \pi/T$) и называется основной.

Выполнение алгебраических действий над ДПЛ не всегда удобно по той причине, что в выражениях для $X(p)$ часто повторяется функция $e^{j\omega T}$. Одним из путей преодоления этого неудобства является введение еще одного вида преобразований над дискретными сигналами — дискретного z -преобразования.

Дискретное z -преобразование является удобным и полезным способом математического представления дискретных сигналов и дискретных систем. Прямое z -преобразование $X(z)$ дискретного сигнала $x(nT)$ получается из прямого ДПЛ введением новой комплексной переменной $z = e^{pT}$:

$$X(z) = Z[x(nT)] = \sum_{n=0}^{\infty} x(nT)z^{-n}, \quad (1.26)$$

где $z = \alpha + j\beta = re^{j\varphi}$. Вследствие выполненной замены выражения для $X(z)$ имеют более простой вид, чем выражения для $X(p)$, что обуславливает широкое применение дискретного z -преобразования.

Комплексная функция $X(z)$, часто называемая z -изображением или z -образом исходного дискретного сигнала, определена лишь для тех значений переменной z , для которых ряд (1.26) является сходящимся, при этом множество соответствующих значений z называется областью сходимости. Условием равномерной сходимости ряда (1.26) является сходимость ряда, состоящего из модулей его членов:

$$|X(z)| = \sum_{n=0}^{\infty} |x(nT)z^{-n}| = \sum_{n=0}^{\infty} |x(nT)|r^{-n} < \infty.$$

Очевидно, что область сходимости в этом случае определяется радиусом R круга в z -плоскости, вне которого ряд (1.26) сходится. На рис. 1.36 область сходимости $X(z)$ показана штриховкой.

Удобным способом графического представления $X(z)$ является изображение нулей z_i^0 и полюсов z_i^* дискретного z -преобразования на z -плоскости, называемое картой нулей и полюсов.

Нулями $X(z)$ называются значения комплексной переменной $z = z_i^0$, $i = 1, 2, 3, \dots$, обращающие выражение (1.26) в нуль, а полюсами — такие значения переменной $z = z_i^*$, $i = 1, 2, 3, \dots$, при которых обеспечивается равенство $X(z) = \infty$.

Пример 1.17. Сигнал представляет собой дискретную экспоненту $x(nT) = a^n$; $a = \text{const}$; $n = 0, 1, 2, \dots$. Требуется рассчитать прямое z -преобразование сигнала.

Z -образ дискретной экспоненты в соответствии с (1.13) определяется следующим образом:

$$Z(a^n) = \sum_{n=0}^{\infty} a^n z^{-n} = \sum_{n=0}^{\infty} (az^{-1})^n = \sum_{n=0}^{\infty} \left(\frac{a}{z}\right)^n.$$

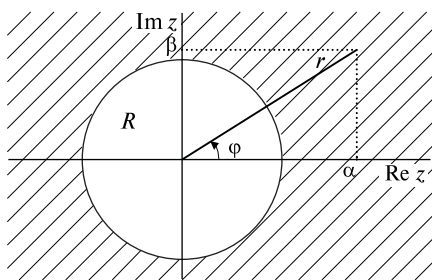


Рис. 1.36. Область сходимости z -преобразования

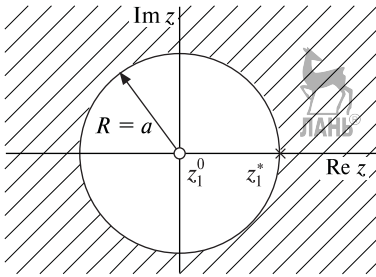


Рис. 1.37. Карта нулей и полюсов, область сходимости z -образа сигнала $x(nT) = a^n$

Для выполнения условия сходимости необходимо, чтобы $|az^{-1}| < 1$, в противном случае сумма ряда будет равна бесконечности. Следовательно, z -образ сигнала $x(nT) = a^n$ определен при условии $|z| > |a|$, т.е. область сходимости определяется кругом радиуса $R = a$, вне которого ряд $Z(a^n)$ сходится и, в соответствии с теорией рядов, имеет значение

$$Z(a^n) = \frac{1}{1 - az^{-1}}.$$

Для определения нулей и полюсов z -образа необходимо преобразовать $Z(a^n)$ к виду без отрицательных степеней переменной:

$$Z(a^n) = \frac{z}{z - a},$$

после чего становится очевидным наличие одного нуля $z_1^0 = 0$ и одного полюса $z_1^* = a$. На рис. 1.37 представлена карта нулей и полюсов и область сходимости z -образа исследуемого сигнала.

Определенный интерес представляют z -образы типовых дискретных сигналов: цифрового единичного импульса и цифрового единичного скачка:

$$Z[\delta(n)] = \sum_{n=0}^{\infty} \delta(n)z^{-n} = \delta(0) \cdot 1 + 0 + 0 + 0 + \dots = 1;$$

$$Z[u_1(n)] = \sum_{n=0}^{\infty} u_1(n)z^{-n} = \frac{1}{1 - z^{-1}} \quad \text{при } |z| > 1.$$

Обратное дискретное z -преобразование ставит в соответствие z -образу $X(z)$ дискретную последовательность $x(nT)$:

$$x(nT) = \frac{1}{2\pi j} \oint_C X(z)z^{n-1}dz,$$

где C — любой замкнутый контур, расположенный в области сходимости $X(z)z^{n-1}$ и охватывающий начало координат z -плоскости.

Ниже представлены основные свойства прямого дискретного z -преобразования.

Свойство линейности. Прямое дискретное z -преобразование линейной комбинации дискретных сигналов равно такой же линейной комбинации z -преобразований этих сигналов.

Если сигнал $x_1(nT)$ имеет z -образ $X_1(z)$, а сигнал $x_2(nT)$ имеет z -образ $X_2(z)$, то сигнал $x_3(nT) = a_1x_1(nT) + a_2x_2(nT)$, где a_1 и a_2 — постоянные коэффициенты, имеет z -изображение $X_3(z) = a_1X_1(z) + a_2X_2(z)$.

Свойство задержки (сдвига). Прямое дискретное z -преобразование дискретного сигнала, сдвинутого относительно исходного на

m отсчетов, равно z -преобразованию исходного сигнала, умноженному на z^{-m} .

Если сигнал $x(nT)$ имеет z -образ $X(z)$, то сигнал $y(nT) = x[(n - m)T]$ имеет z -образ $Y(z) = X(z)z^{-m}$.

Свойство z -преобразования линейной (апериодической) свертки. Прямое дискретное z -преобразование дискретного сигнала, равного апериодической свертке двух исходных сигналов, равно произведению z -преобразований исходных сигналов.

Если сигнал $x_1(nT)$ имеет z -образ $X_1(z)$, а сигнал $x_2(nT)$ — z -образ $X_2(z)$, то сигнал

$$y(nT) = \sum_{m=0}^n x_1(mT)x_2(nT - mT) = \sum_{m=0}^n x_1(nT - mT)x_2(mT)$$

имеет z -изображение $Y(z) = X_1(z)X_2(z)$.

1.16. Связь между преобразованиями Лапласа, Фурье и z -преобразованием дискретных сигналов

Равноправность использования в теории ЦОС преобразований Лапласа, Фурье и z -преобразования требует знания связей между ними, которые можно выявить на основе сравнения выражений для прямых преобразований:

$$X(p) = \sum_{n=0}^{\infty} x(nT)e^{-pnT};$$

$$X(e^{j\omega T}) = \sum_{n=0}^{\infty} x(nT)e^{-j\omega nT};$$

$$X(z) = \sum_{n=0}^{\infty} x(nT)z^{-n}.$$

Сравнение выражений (1.25) и (1.5) показывает, что, как и в случае с непрерывными сигналами, ДПТ является обобщением преобразования Фурье, распространяя его со случая чисто мнимого аргумента $j\omega$ на более общий случай комплексного аргумента $p = \sigma + j\omega$.

Сравнение преобразования Фурье и дискретного z -преобразования (выражения (1.5) и (1.26)) свидетельствует о том, что при условии сходимости рядов спектр дискретного сигнала может быть получен путем замены переменной в z -образе этого сигнала в соответствии с равенством $z = e^{j\omega T}$. Данное уравнение отображает точки z -плоскости, расположенные на окружности радиусом $r = 1$ с центром в начале координат. Следовательно, спектр дискретного сигнала соответствует его z -образу, рассматриваемому на единичной

окружности в z -плоскости:

$$X(e^{j\omega T}) = X(z)|_{z=e^{j\omega T}},$$

а ДПФ периодического дискретного сигнала представляет собой z -преобразование, определенное в N равноотстоящих точках этой окружности, соответствующих угловым частотам $k\Omega = k2\pi/(NT)$, $k = 0, 1, 2, \dots, N-1$.

Выяснение связи между ДПЛ и z -преобразованием достаточно полно характеризуется соответствием точек p - и z -плоскостей. Известно, что

$$z = e^{pT} = e^{(\sigma + j\omega)T} = e^{\sigma T} e^{j\omega T} = \alpha + j\beta = re^{j\varphi},$$

тогда координаты любой точки z -плоскости выражаются через координаты соответствующей точки p -плоскости следующим образом: $r = e^{\sigma T}$; $\varphi = \omega T + 2\pi k$; $\alpha = e^{\sigma T} \cos \omega T$; $\beta = e^{\sigma T} \sin \omega T$. Графическая иллюстрация полученных выражений для основной полосы ДПЛ показана на рис. 1.38.

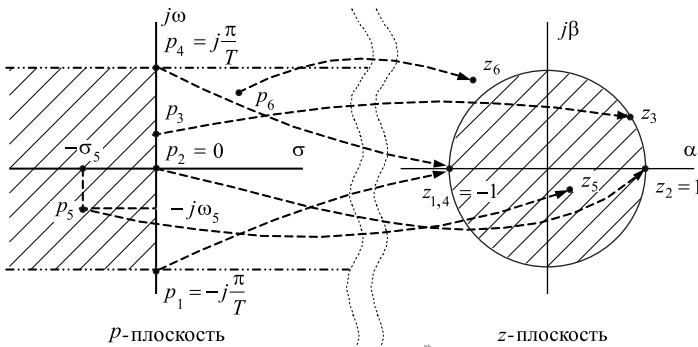


Рис. 1.38. Соответствие между точками p - и z -плоскостей в основной полосе ДПЛ

При движении точки p по мнимой оси ($\sigma = 0$) от $p = -j\infty$ до $p = j\infty$ соответствующая ей точка $z = e^{j\omega T}$ последовательно описывает в z -плоскости окружности единичного радиуса ($|z| = 1$) по направлению против часовой стрелки, т.е. мнимая ось p -плоскости отображается в единичную окружность z -плоскости. При этом точки $p = j(\pi/T \pm 2k\pi/T)$; $k = 0, 1, 2, \dots$, отображаются в точку $z = -1$, а точка начала координат p -плоскости ($p_1 = 0$) и точки $p = \pm j \cdot 2k\pi/T$ переходят в точку $z_1 = 1$ на вещественной оси z -плоскости.

Очевидно, что внутри любой полосы ДПЛ шириной $2\pi/T$ точки левой p -полуплоскости ($\sigma < 0$) отображаются внутрь единичной окружности z -плоскости, а точки правой p -полуплоскости ($\sigma > 0$) —

на всю z -плоскость, исключая единичный круг. Точке начала координат в z -плоскости ($z = 0$) соответствует точка $p = -\infty$.

Таким образом, z -плоскость не загромождена повторяющейся информацией различных полос дискретного преобразования Лапласа.

1.17. Вейвлет-преобразование

1.17.1. Особенности применения и использования вейвлет-анализа

В настоящее время одним из основных средств анализа реальных физических процессов является гармонический анализ на основе преобразования Фурье, обладающий рядом недостатков:

1. Гармонические базисные функции разложения не способны отображать перепады сигналов с бесконечной крутизной типа прямоугольных импульсов, так как это требует бесконечно большого числа членов ряда. При ограничении числа членов ряда Фурье возникают осцилляции (явление Гиббса) в окрестностях скачков и разрывов восстановленного сигнала.

2. Особенности сигналов, связанные с разрывами (скачками) и острыми пиками, вызывают изменения их частотного образа, что делает невозможным их обнаружение и анализ по спектральному представлению данных сигналов.

3. Преобразованием Фурье отображаются глобальные сведения о частотах исследуемого сигнала, что не дает представления о локальных свойствах сигнала при быстрых временных изменениях его спектрального состава.

Если во временной области сигнал состоит из нескольких последовательно с течением времени сменяющих друг друга различных гармонических сигналов, то выполнение преобразования Фурье над таким сигналом приведет к спектральному представлению, в котором ошибочно подразумевается, что спектральные компоненты существовали на всем интервале анализа сигнала. Таким образом, преобразование Фурье в принципе не имеет возможности анализировать частотные характеристики сигнала в произвольные моменты времени.

Решить данную проблему можно с помощью уменьшения временного интервала анализа, применив допущение о малом размере участка стационарности анализируемого сигнала. Такой подход получил название оконного (или кратковременного) преобразования Фурье. Сигнал делится на отрезки — «окна», в пределах которых его можно считать стационарным, для чего на сигнал накладывается оконная функция. Однако при таком подходе существует ряд

принципиальных проблем, связанных с принципом неопределенности Гейзенберга. Согласно этому принципу, невозможно получить произвольно точное частотно-временное представление сигнала, т. е. нельзя определить, какие спектральные составляющие присутствуют в сигнале, если его длительность бесконечно мала. Эта проблема называется проблемой разрешения.

Более подробно методы спектрального анализа с использованием оконных функций будут представлены в главе 4.

Вейвлет-анализ позволяет перейти к более универсальному частотно-временному представлению сигнала. Первые теоретические работы по основам вейвлет-преобразований были выполнены в 90-х годах прошлого века Мейером (Meyer Y.), Добеши (Daubechies I.) и Маллатом (Mallat S.A.). Математический аппарат вейвлет-преобразований находится в стадии активной разработки, однако специальные пакеты расширений по вейвлетам уже присутствуют в основных системах компьютерной математики (MATLAB, Mathcad, Mathematica и др.).

Теория вейвлетов не является фундаментальной физической теорией, но она дает удобный инструмент для решения многих практических задач. Основная область применения вейвлет-преобразований — анализ и обработка сигналов и функций, нестационарных во времени или неоднородных в пространстве, когда результаты анализа должны содержать не только общую частотную характеристику сигнала (распределение энергии сигнала по частотным составляющим), но и сведения об определенных временных координатах, на которых себя проявляют те или иные группы частотных составляющих, или на которых происходят быстрые изменения частотных составляющих сигнала. По сравнению с разложением сигналов на ряды Фурье, вейвлеты способны с гораздо более высокой точностью представлять локальные особенности сигналов, вплоть до разрывов огибающей 1-го и 2-го рода. В отличие от преобразований Фурье, вейвлет-преобразование одномерных сигналов обеспечивает двумерную развертку, при этом частота и координата рассматриваются как независимые переменные, что дает возможность анализа сигналов сразу в двух пространствах.

Из практики применения преобразования Фурье известно, что используется окно конкретной длительности для анализа всего сигнала, тогда как разные его участки могут требовать применения различных окон. Если сигнал состоит из далеко отстоящих друг от друга частотных компонентов, то можно пожертвовать спектральным разрешением в пользу временного и наоборот. Операция умножения на окно как бы содержится в самой базисной функции вейвлет-

анализа, которая позволяет сужать и расширять это окно. Отсюда появляется возможность адаптивного к сигналу выбора параметров окна.

Одна из главных идей вейвлет-представления сигналов на различных уровнях декомпозиции (разложения) заключается в разделении функций приближения к сигналу на две группы: аппроксимирующую — грубую, с достаточно медленной временной динамикой изменений, и детализирующую — с локальной и быстрой динамикой изменений на фоне плавного изменения, с последующим их дроблением и детализацией на других уровнях декомпозиции сигналов. Это возможно как во временной, так и в частотной областях представления сигналов вейвлетами.

Особенно важна принципиальная возможность вейвлетов представлять нестационарные сигналы, состоящие, например, из различных компонентов, действующих на непересекающихся временных интервалах, модулированные сигналы и т. д. Вейвлет-спектрограммы значительно информативнее обычных спектрограмм, что делает их незаменимыми при «тонком» анализе структуры сигналов в различных прикладных задачах, например при идентификации параметров цифровых систем.

Применение вейвлетов в задачах анализа и кодирования изображений продиктовано особенностями его представления. Вейвлеты, как средство многомасштабного анализа, позволяют выделять одновременно как основные характеристики сигнала, так и короткоживущие высокочастотные явления в изображении. Это свойство является существенным преимуществом в задачах обработки сигнала по сравнению с «оконным» базовым преобразованием Фурье, где, варьируя длительность сегмента анализа и используемые оконные функции, приходится выбирать масштаб явлений, которые необходимо выделить в сигнале. Получение дополнительной информации с разных масштабов времени и разных масштабов разрешения сигнала улучшает точность анализа изображения.

Вейвлеты (wavelet — короткая волна, иногда переводится как всплеск) — функции определенной формы, локализованные по оси аргументов (независимых переменных), инвариантные к сдвигу и линейные к операции масштабирования (сжатия/растяжения).

Гармонические базисные функции преобразования Фурье предельно локализованы в частотной области (до импульсных функций Дирака при) и не локализованы во временной (определены во всем временном интервале от до). Их противоположностью являются импульсные базисные функции типа импульсов Кронекера, которые предельно локализованы во временной области и «размыты» по все-

му частотному диапазону. Вейвлеты по локализации в этих двух представлениях можно рассматривать как функции, занимающие промежуточное положение между гармоническими и импульсными функциями. Они должны быть локализованными как во временной, так и в частотной областях представления. Особенностью вейвлет-анализа является то, что в нем можно использовать семейства функций, реализующих различные варианты соотношения неопределенности. Исследователь имеет возможность гибкого выбора между ними и применения тех вейвлет-функций, которые наиболее эффективно решают поставленные задачи.

1.17.2. Непрерывный вейвлет-анализ

Вейвлет-базис пространства $L^2(R)$ (Гильберта), $R(-\infty, \infty)$ целесообразно конструировать из семейства финитных (компактных) функций, принадлежащих этому же пространству, которые должны стремиться к нулю на бесконечности. Чем быстрее эти функции стремятся к нулю, тем удобнее использовать их в качестве базиса преобразования при анализе реальных сигналов.

Функции вейвлетов создаются на основе базисной функции $\psi_0(t)$, определяющий тип вейвлета и удовлетворяющей свойствам смещения по оси времени и масштабирования. Оба свойства объединяются выражением

$$\psi_{ab}(t) = \frac{1}{\sqrt{|a|}} \psi_0\left(\frac{t-b}{a}\right).$$

Числовой параметр a задает ширину вейвлета, а числовой параметр b — его положение на оси времени.

В общем случае параметры a и b могут меняться непрерывно, что в итоге определит бесконечное множество возможных базисных функций. Необходимо ограничить мощность множеств значений a и b при сохранении возможности восстановления сигнала из его преобразования. Для этого необходимо ввести параметры масштаба m и сдвига k , приняв

$$a = a_0^m; \quad b = b_0 k a_0^m, \quad (1.27)$$

где $m, k \in Z$ (область целых чисел); $a_0 > 1$ — исходный масштаб материнского вейвлета; b_0 — положение на оси времени материнского вейвлета.

Из выражения (1.27) видно, что параметр местоположения зависит от параметра масштаба, с увеличением которого увеличивается размер шага сдвига. Это объясняется тем, что при анализе с большим масштабом детали становятся менее важны. Таким образом,

определяется семейство *psi*-функций, имеющих относительно конечную мощность параметров масштаба и сдвига:

$$\psi_{mk}(t) = a_0^{-m/2} \psi_0(a_0^{-m}t - k).$$

Данные функции равны нулю за пределами некоторого конечного интервала, если для них выполняется условие ортогональности

$$\langle \psi_{mk}(t), \psi_{lm}(t) \rangle = \int_{-\infty}^{\infty} \psi_{mk}(t) \psi_{lm}^*(t) dt = \delta_{nl} \delta_{km}.$$

В качестве примера графика такой функции рассмотрим *psi*-функцию вейвлета, носящего название «мексиканская шляпа», при $t \in (-5; 5)$ (рис. 1.39):

$$\psi(t) = \left(\frac{2}{\sqrt{3}} \pi^{-1/4} \right) (1 - t^2) e^{-t^2/2}.$$

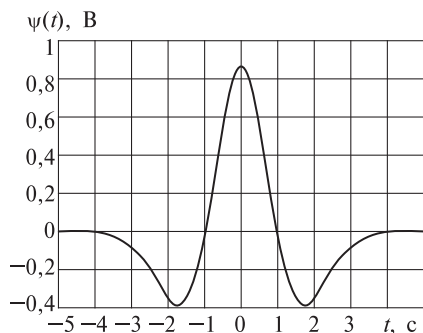


Рис. 1.39. График *psi*-функции вейвлета «мексиканская шляпа»

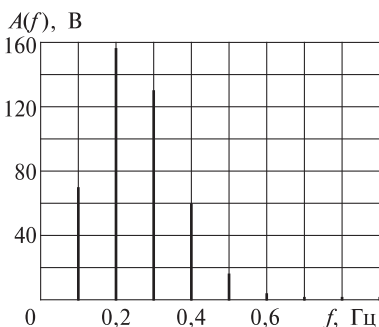


Рис. 1.40. Амплитудный спектр вейвлета «мексиканская шляпа»

Вейвлет может применяться для анализа сложных сигналов, он симметричен, его спектр представлен только действительной частью и хорошо локализован по частоте.

Отсчеты амплитудного спектра данного вейвлета, полученный с использованием ДПФ, представлен на рис. 1.40 (на нем показаны первые 10 спектральных составляющих, частота дискретизации исходного вейвлета равна 100 Гц).

Постоянная составляющая спектра, а также все последующие спектральные составляющие, начиная с восьмой, имеют нулевые значения, что подтверждает локализованность вейвлета в спектральной области.

Из вышесказанного следует, что произвольная функция $S(t)$ пространства $L^2(R)$ может быть представлена в виде ряда (разложения по базису), где коэффициенты представления сигнала — проекции сигнала на новый ортонормированный базис функций, как и

в преобразовании Фурье, определяются скалярным произведением

$$S_{mk} = \langle s(t), \psi_{mk}(t) \rangle = \int_{-\infty}^{\infty} s(t) \psi_{mk}(t) dt, \quad (1.28)$$

при этом ряд равномерно сходится:

$$\lim_{M, K \rightarrow \infty} \left\| s(t) - \sum_{m=-M}^M \sum_{k=-K}^K S_{mk} \psi_{mk}(t) \right\| = 0.$$

Для случая применения ортогональных вейвлетов выражение (1.28) преобразуется к виду

$$S_{mk} = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} s(t) \psi_{mk}(t) dt.$$

Вейвлет-спектр, в отличие от преобразования Фурье, является двумерным и определяет поверхность в пространстве переменных m и k . Особенностью являются две оси независимых переменных — ось сдвига (времени), ось масштаба (частоты) и третья ось, определяющая значения функции вейвлет-разложения.

В практических задачах достаточны устойчивость выполняемых преобразований и «приблизительность» ортогональности системы базисных функций разложения. Под устойчивостью понимается достаточно точная реконструкция произвольных сигналов. Для ортогональных вейвлетов обратное вейвлет-преобразование записывается с помощью того же базиса, что и прямое:

$$s(t) = \frac{1}{C_\psi} \int_R \frac{1}{a^2} c(m, k) \psi_{mk}(t) dm dk,$$

где C_ψ — нормализующий коэффициент:

$$C_\psi = \int_{-\infty}^{\infty} \frac{|\Psi(\omega)|^2}{\omega} d\omega < \infty. \quad (1.29)$$

Условие конечности C_ψ ограничивает класс функций, которые можно использовать в качестве вейвлетов. В частности, при $\omega = 0$ для обеспечения сходимости интеграла (1.29) значение $\Psi(\omega)$ должно быть равно нулю, а функция $\psi(t)$ должна удовлетворять следующему выражению:

$$\int_{-\infty}^{\infty} \psi(t) dt = 0. \quad (1.30)$$

В частотной области требование о нулевом значении интеграла функции $\psi(t)$ указывает на смещение частотного образа физически реализуемого вейвлета по оси частот с сосредоточением вокруг центральной частоты f_0 .

Однако это означает, что не для всех сигналов возможна их точная аппроксимация вейвлетом $\psi_{mk}(t)$, так как при выполнении (1.30) вейвлет-коэффициент при постоянной составляющей сигнала равен нулю. Кроме того, далеко не все типы вейвлетов могут гарантировать реконструкцию сигналов как таковую. Однако и такие вейвлеты могут быть полезны для анализа особенностей сигналов как дополнительного метода к другим методам анализа и обработки данных.

Вейвлет-преобразование обладает рядом свойств, отличающих его от других преобразований и определяющих его достоинства и недостатки. Рассмотрим основные из этих свойств. Для обозначения операции вейвлет-преобразования произвольной функции $s(t)$ применим форму записи $VP[s(t)]$.

Основные свойства вейвлет преобразования:

- линейность

$$VP[\alpha s_1(t) + \beta s_2(t)] = \alpha VP[s_1(t)] + \beta VP[s_2(t)];$$

- инвариантность относительно сдвига

$$VP[s(t - t_0)] = C(m, k - t_0);$$

- инвариантность относительно масштабирования

$$VP[s(t/m_0)] = (1/m_0)C(m/m_0, k/a_0);$$

- дифференцирование

$$\frac{d\{VP[s(t)]\}}{dt} = VP\left[\frac{d(s(t))}{dt}\right];$$

$$VP\left[\frac{d^n(s(t))}{dt^n}\right] = (-1)^n \int_{-\infty}^{\infty} s(t) \frac{d^n(\psi(t))}{dt^n} dt;$$

- свойство сохранения энергии

$$\int_R s_1(t) s_2^*(t) = C_{\psi}^{-1} \int_R a^{-2} C(m, k) C^*(m, k) dm dk.$$

Достоинства и недостатки вейвлет-преобразований:

- они обладают всеми достоинствами преобразований Фурье;
- вейвлет-базисы могут быть хорошо локализованными как по частоте, так и по времени. При выделении в сигналах хорошо локализованных разномасштабных процессов можно рассматривать только уровни разложения, представляющие целевой интерес;
- вейвлет-базисы, в отличие от преобразования Фурье, имеют достаточно много разнообразных базовых функций, свойства которых ориентированы на решение различных задач;
- относительная сложность.

1.17.3. Дискретный вейвлет-анализ

Одним из существенных преимуществ вейвлет-преобразования перед преобразованием Фурье является то, что для анализа различных сигналов могут применяться различные вейвлет- и скейлинг-функции (scale — масштаб), выделяющие особенности сигнала. Следует отметить, что не всегда ставится задача восстановления сигнала по его вейвлет-спектру, а следовательно, множество применяемых к сигналу вейвлет- и скейлинг-функций существенно расширяется.

Для практического применения преобразования необходимо перейти от непрерывного вейвлет-преобразования к его дискретному виду, что позволяет получать быстрые алгоритмы расчета для внедрения в цифровые средства обработки информации.

При применении вейвлет-анализа используются функции $s(t)$ с конечной энергией (нормой) в пространстве $L^2(R)$, определенные по всей действительной оси $R(-\infty, \infty)$. Для финитных сигналов с конечной энергией средние значения сигналов, как и любых других функций из пространства $L^2(R)$, должны стремиться к нулю на $\pm\infty$, следовательно, для разложения сигнала $s(nT)$ справедливо

$$s(nT) = \sum_m a_m \varphi_m(nT) + \sum_k d_k \psi_k(nT),$$

где $\varphi_m(nT)$ — масштабирующие (скейлинг) функции (phi-функции) с единичным значением суммы отсчетов $\sum_{-\infty}^{\infty} \varphi_m(nT) = 1$, которые определяют грубое приближение сигнала и порождающие коэффициенты аппроксимации a_m ; phi-функции присущи не всем вейвлетам и необходимы для преобразования нецентрированных и достаточно протяженных сигналов при раздельном анализе низкочастотных и высокочастотных составляющих, что чаще всего и встречается на практике;

$\psi_k(nT)$ — вейвлет-функции (psi-функции) с нулевым значением суммы отсчетов $\sum_{-\infty}^{\infty} \psi(nT) = 0$. Этой функцией, которую обычно и называют вейвлетом, выделяются детали сигнала и его локальные особенности, определяющие детализированность сигнала и порождающие коэффициенты d_k .

Если анализируемый сигнал является дискретным, то масштабирующие и вейвлет-функции заменяются значениями их отсчетов во временной области, при этом, как правило, используется процедура кратномасштабного анализа. Значения отсчетов соответствующего вейвлета по своей сути определяют импульсные характеристики ФНЧ-фильтра (аппроксимирующей phi-функции) и ФВЧ-фильтра

(детализирующей psi -функции) КИХ-фильтров. Более подробно о данных типах фильтров будет изложено в главе 2. При таком представлении существует возможность как разложения, так и восстановления сигнала. При использовании вейвлет для получения импульсных характеристик фильтров разложения сигнал восстанавливается с любого этапа декомпозиции сигнала.

Произвольный дискретный сигнал обычно представляет собой сумму разнотипных составляющих: медленно меняющихся функций (средних значений по большим интервалам усреднения), циклических компонентов с определенным периодом повторения достаточно гладких по форме и быстро меняющихся функций, отражающих локальные особенности и резкие изменения сигнала в определенные моменты времени. Инструментом разделения сигналов на такие составляющие и анализа их порядка является кратномасштабный анализ. Исходный сигнал раскладывается на совокупность низкочастотных и высокочастотных составляющих, при этом на каждом этапе может осуществляться децимация временной последовательности составляющих спектра с коэффициентом $M = 2$, о процедуре децимации более подробно будет изложено в главе 3.

Можно заметить, что в результате первого шага разложения временное разрешение уменьшается в два раза, так как лишь половина отсчетов характеризует весь сигнал. Однако частотное разрешение удваивается, так как сигнал занимает теперь половинную полосу частот и неопределенность уменьшается. Далее сигнал с выхода низкочастотного фильтра поступает на такую же схему обработки, а сигнал на выходе высокочастотного фильтра считается вейвлет-коэффициентами. Сигнал на выходе высокочастотного фильтра также может быть подвержен децимации. Этот процесс показан на рис. 1.41, где $x(n)$ — исходный сигнал, $h_{\text{НЧ}}(n)$ и $h_{\text{ВЧ}}(n)$ — импульсные характеристики низкочастотного и высокочастотного фильтров соответственно. Спектральная область сигнала на каждом уровне обозначена как « ω_0 ».

Используя рис. 1.41, рассмотрим пример. Предположим, что сигнал $x(n)$ имеет 512 отсчетов и частотную полосу от 0 до π рад. После первого уровня декомпозиции на выходе высокочастотного фильтра после децимации имеется 256 отсчетов. Разрешение по времени уменьшилось в два раза, но частотный диапазон, занимаемый сигналом, теперь от $\pi/2$ до π рад/с (т. е. разрешение по частоте удвоилось). Эти 256 отсчетов есть вейвлет-коэффициенты первого уровня. Сигнал на выходе низкочастотного фильтра после децимации занимает полосу частот от 0 до $\pi/2$ рад/с. Этот сигнал поступает опять на два фильтра: низкочастотный и высокочастотный.

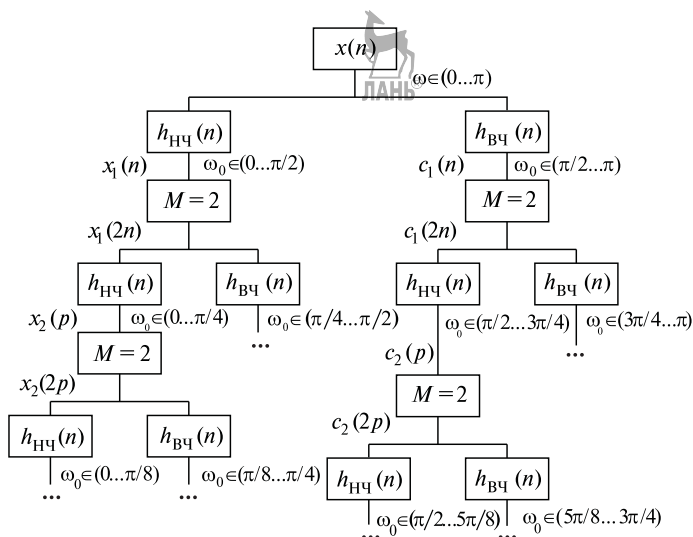


Рис. 1.41. Разложение сигнала при кратномасштабном вейвлет-анализе

Этот процесс может продолжаться до тех пор, пока не останутся 2 отсчета, т.е. для данной длины сигнала может быть 8 уровней декомпозиции. В итоге дискретное вейвлет-преобразование исходного сигнала получается путем объединения вейвлет-коэффициентов всех уровней, начиная с последнего (2 коэффициента). К ним присоединяются спереди отсчеты аппроксимации сигнала последнего уровня (2 отсчета). Таким образом, число коэффициентов преобразования равно числу отсчетов в исходном сигнале.

Наиболее значимые частоты исходного сигнала будут отображаться как большие амплитуды вейвлет-коэффициентов, «ответвляющих» за соответствующий частотный диапазон. Отличие от ДПФ заключается в том, что время появления частот в данном случае не утеряно. Однако временная локализация будет иметь разрешение, зависящее от уровня преобразования, на котором появляется частота. Если основная информация сигнала лежит в области высоких частот, как это часто случается, то временная локализация этих частот будет более точной, так как они характеризуются большим количеством отсчетов. На низких частотах, напротив, отсчетов мало, поэтому временное разрешение будет плохим, зато хорошим будет частотное разрешение (так как сигнал занимает узкую полосу частот). Такая схема анализа подходит для большинства сигналов, встречающихся на практике. Малые значения вейвлет-коэффициентов означают низкую энергетику соответствующих частотных полос в сигнале. Эти коэффициенты могут быть отброшены без существенного

искажения сигнала. Таким образом, может достигаться уменьшение данных (сжатие). Процедуру обнуления незначущих коэффициентов вейвлет-разложения называют трешолдингом, она может применяться при эффективном кодировании данных с потерями и при реализации систем шумоподавления.

Важным свойством такого подхода является взаимосвязь между импульсными характеристиками высокочастотного и низкочастотного фильтров.

Эти фильтры связаны между собой отношением

$$h_{\text{ВЧ}}[L - 1 - n] = (-1)^n - h_{\text{НЧ}}(n),$$

где L — порядок фильтра.

Таким образом, высокочастотный фильтр получается из низкочастотного путем «переворота» вектора его коэффициентов и изменения знака у нечетных коэффициентов. Операции фильтрации и прореживания могут быть записаны в виде:

$$c[k] = \sum_n x(n) h_{\text{ВЧ}}[-n + 2k]; \quad y_{\text{low}}[k] = \sum_n x(n) h[-n + 2k].$$

Синтез выполняется в обратном порядке по сравнению с анализом. Сигналы на каждом уровне интерполируются в 2 раза, пропускаются через фильтры синтеза и затем складываются, фильтры анализа и синтеза идентичны, за исключением обратного порядка следования коэффициентов. Выражение для восстановления сигналов во временной области выглядит следующим образом:

$$x[n] = \sum_{k=-\infty}^{\infty} c(k) h[-n + 2k].$$



2 Основы цифровой фильтрации



2.1. Дискретные системы

В научной литературе приводится множество определений системы, но применительно к задачам анализа и синтеза информационно-телекоммуникационных систем и систем управления полезным будет следующее. *Системой* называется упорядоченная совокупность элементов (звеньев, устройств), связанных друг с другом направленными связями, действующих как одно целое и предназначенных для достижения какой-либо общей цели. В ЦОС рассматриваются *дискретные системы* (ДС) — системы, имеющие дискретную временную шкалу и обрабатывающие дискретные сигналы.

Для описания свойств систем широко используются математические модели и, в частности, аппарат функционального анализа (теория множеств). Пусть $\{X\}, \{Y\}$ — пространства (множества) дискретных входных и выходных сигналов. Если при этом каждому сигналу $x(nT) \in X$ соответствует по определенному закону сигнал $y(nT) \in Y$, то говорят, что задан оператор системы F . Нахождение оператора системы F можно назвать также процессом построения математической модели системы. Понятия пространства сигналов, оператора системы позволяют применять строгие математические методы для анализа и синтеза систем, что особенно важно в период массового применения компьютеров. На рис. 2.1 показана теоретико-множественная модель дискретной системы.

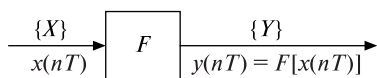


Рис. 2.1. Теоретико-множественная модель дискретной системы

Под математической моделью системы следует понимать количественную формализацию абстрактных представлений об изучаемой системе. *Математическая модель* —

это формальное описание системы с помощью математических средств: дифференциальных, интегральных и алгебраических уравнений, неравенств, множеств.

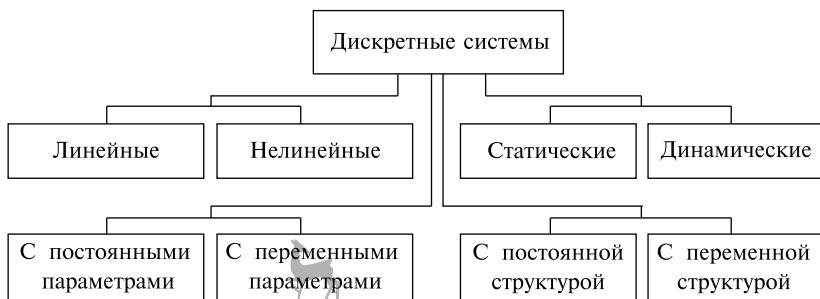


Рис. 2.2. Вариант классификации дискретных систем

На рис. 2.2 представлен вариант классификации дискретных систем в зависимости от свойств оператора F .

Дискретная система, в которой значение выходного сигнала $y(nT)$ в момент времени nT определяется значением входного сигнала $x(nT)$ в тот же момент времени, называется *статической системой* или системой «без памяти». Дискретная система, в которой значение выходного сигнала $y(nT)$ в момент времени nT определяется значениями входного сигнала в моменты времени, предшествующие моменту времени nT , называется *динамической системой* или системой «с памятью».

Если для оператора F выполняются условия однородности:

$$F\{ax(nT)\} = aF\{x(nT)\}$$

и аддитивности:

$$F\{x_1(nT) + x_2(nT)\} = F\{x_1(nT)\} + F\{x_2(nT)\},$$

где a — постоянный коэффициент, то такая дискретная система называется *линейной*, в противном случае — *нелинейной*. Принцип суперпозиции, по выполнению которого системы можно разделить на линейные и нелинейные, можно также сформулировать в следующем виде: если входным последовательностям $x_1(nT)$ и $x_2(nT)$ соответствуют реакции (выходные последовательности) $y_1(nT)$ и $y_2(nT)$ соответственно, то входной последовательности $x(nT) = ax_1(nT) + bx_2(nT)$ будет соответствовать реакция $y(nT) = ay_1(nT) + by_2(nT)$.

Если система имеет постоянные параметры, а входной последовательности $x(nT)$ соответствует реакция $y(nT)$, то задержанной на m периодов дискретизации входной последовательности $x(nT - mT)$ соответствует реакция $y(nT - mT)$.

Самым распространенным классом дискретных систем, используемых в теории и практике ЦОС, являются линейные дискретные динамические системы (ЛДДС) с постоянными параметрами и

структурой, которые в дальнейшем будут рассматриваться как линейные цифровые фильтры с постоянной структурой и параметрами.

2.2. Принцип цифровой фильтрации

Фильтрация сигналов является одной из важных задач обработки сигналов в системах телекоммуникаций. Самое простое определение фильтра может быть сформулировано следующим образом: фильтром называется устройство, предназначенное для отделения одной части входной совокупности сигналов и помех (входного сигнала) от другой ее (его) части. Иначе говоря, радиотехническая фильтрация изменяет некоторые спектральные составляющие обрабатываемого сигнала (шумов, помех) при минимальных искажениях полезных спектральных составляющих в определенной полосе частот, таким образом, в общем случае фильтрация — это обработка сигнала, приводящая к изменению значений его спектральных составляющих.

Достаточно широкое применение в реализации задачи фильтрации находят дискретные системы, способные, в общем случае, осуществлять различные преобразования электрических сигналов. Класс дискретных систем, предназначенных для выполнения процедуры фильтрации дискретных сигналов, получил название цифровых (дискретных) фильтров. Таким образом, *цифровая фильтрация* является частным случаем радиотехнической фильтрации, характеризующимся использованием дискретных (цифровых) сигналов и фильтров.

Цифровой фильтр (ЦФ) представляет собой дискретную систему, реализующую преобразование множества входных дискретных сигналов X во множество выходных сигналов Y в соответствии с оператором фильтрации $F: X \rightarrow Y$. Конструктивно ЦФ представляет собой соединение элементов цифровой техники — счетчиков, регистров сдвига, умножителей (аппаратная реализация) или программу для ЭВМ (программная реализация). В обоих случаях под термином «цифровой фильтр» следует понимать совокупность вычислителя, в качестве которого выступает специализированное устройство или ЭВМ, и алгоритма, реализованного конструктивно или с помощью специальной программы и обеспечивающего вычислительную процедуру в соответствии с требуемым оператором F .

К достоинствам аппаратной реализации ЦФ следует отнести повышенное быстродействие устройства, его низкие стоимость и массогабаритные показатели. Основным недостатком такого способа реализации является недостаточная гибкость устройства, препятствующая его использованию в иных целях. Прямо противоположные

достоинства и недостатки, что очевидно, имеет программная реализация цифровых фильтров.

Цифровые фильтры могут использоваться и для фильтрации аналоговых сигналов, что требует предварительного выполнения процедуры аналого-цифрового преобразования с обеспечением возможности последующего перехода к непрерывной форме отфильтрованной последовательности. Такой подход позволяет повысить эффективность процедуры фильтрации, что обусловлено некоторыми преимуществами ЦФ по сравнению с аналоговыми фильтрами. К числу основных преимуществ относятся:

- высокая стабильность характеристик ЦФ, обеспечиваемая уменьшенной зависимостью параметров элементной базы от внешних условий;
- возможность реализации фильтрующих характеристик, трудно достигаемых (а зачастую — недостижимых) при использовании аналоговых фильтров;
- гибкость управления цифровыми фильтрами в случае их программной реализации;
- повышенная надежность работы цифровых устройств;
- уменьшенная масса и габаритные размеры;
- унифицированное конструктивное исполнение.

С другой стороны, переход к цифровой фильтрации аналоговых сигналов сопровождается известными издержками, обусловленными реализацией процедур аналого-цифрового и цифро-аналогового преобразований (введение шумов квантования, дополнительные стоимостные и аппаратные затраты) и общими особенностями ЦОС (наличие шумов округления и т. д.).

2.3. Математические модели и характеристики цифровых фильтров

Входная $x(nT)$ и выходная $y(nT)$ последовательности цифрового фильтра с постоянными параметрами связаны *линейным разностным уравнением* (ЛРУ) вида

$$\sum_{m=0}^{M-1} a_m y(nT - mT) = \sum_{k=0}^{K-1} b_k x(nT - kT), \quad (2.1)$$

где M и K — постоянные целые числа, a_m и b_k — постоянные коэффициенты; n — нормированное дискретное время. Название «разностное уравнение» вытекает из понятия «конечная разность», которое соответствует понятию «производная» в теории непрерывных систем управления. Конечные разности так же, как и производные,

имеют различный порядок. Оператор конечной разности i -го порядка D_i соответствует оператору дифференцирования d^i/dt^i , что отражается в соответствующих выражениях:

$$D_0[x(nT)] = x(nT);$$

$$D_1[x(nT)] = x(nT) - x(nT - T);$$

$$D_2[x(nT)] = D_1[x(nT) - x(nT - T)] = x(nT) - 2x(nT - T) + x(nT - 2T).$$

Если $a_0 = 1$, то

$$y(nT) = - \sum_{m=1}^{M-1} a_m y(nT - mT) + \sum_{k=0}^{K-1} b_k x(nT - kT). \quad (2.2)$$

Выражения (2.1) и (2.2) представляют собой рекурсивные ЛРУ с постоянными коэффициентами. Цифровой фильтр, соответствующий такому ЛРУ, является *рекурсивным* (рис. 2.3). Рекурсивный характер уравнений определяется тем, что в формировании текущего отсчета выходного сигнала участвуют не только отсчеты входного сигнала, но и предыдущие отсчеты реакции.

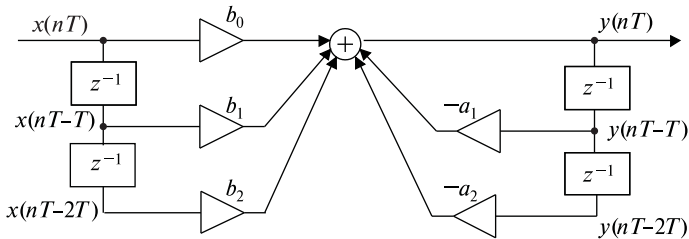


Рис. 2.3. Рекурсивный цифровой фильтр

Если при $k = 1, 2, 3, \dots, K - 1$ коэффициенты $b_k = 0$, то уравнение (2.2) принимает вид

$$y(nT) = b_0 x(nT) - \sum_{m=1}^{M-1} a_m y(nT - mT), \quad (2.3)$$

который определяет рекурсивный характер дискретной системы.

Из уравнения (2.2) видно, что выходной сигнал $y(nT)$ в момент времени nT определяется алгебраической суммой взвешенных значений входного сигнала $x(nT)$ в моменты времени $nT, nT - 1T, nT - 2T$ и т. д., а также взвешенных сигналов выходной последовательности $y(nT)$ в моменты времени $nT - 1T, nT - 2T, \dots$. Следовательно, для вычисления реакции фильтра в момент $nT = 0$ необходимо знать

значения $y(-T)$, $y(-2T)$, $y(-3T)$ и т. д., т. е. значения реакции системы в моменты времени, предшествующие началу анализа. Состояние цифрового фильтра в момент $nT = 0$ характеризуется *начальными условиями*. Если при входном воздействии $x(0) = 0$ выходной сигнал $y(0) = 0$, то говорят о *нулевых начальных условиях*.

Если при $m = 1, 2, 3, \dots, M - 1$ коэффициенты $a_m = 0$, то уравнение (2.2) принимает вид

$$y(nT) = \sum_{k=0}^{K-1} b_k x(nT - kT). \quad (2.4)$$

Выходной сигнал $y(nT)$ фильтра, определяемый выражением (2.3), формируется алгебраической суммой взвешенных входных сигналов $x(nT - kT)$ и не зависит от предшествующих значений реакции системы. Такой ЦФ, в отличие от рекурсивного, описываемого формулой (2.2), называется *нерекурсивным* (рис. 2.4).

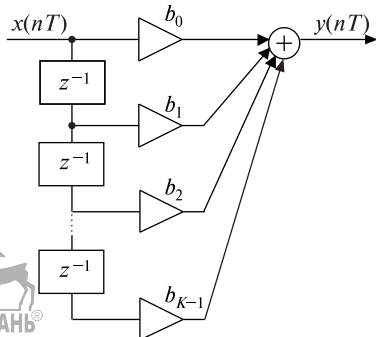


Рис. 2.4. Нерекурсивный цифровой фильтр

Таким образом, уравнения (2.2) и (2.4) являются математическими моделями рекурсивного и нерекурсивного цифровых фильтров соответственно в форме линейных разностных уравнений с постоянными коэффициентами.

Математические модели ЦФ позволяют проводить анализ линейных цифровых фильтров. Из уравнений видно, что весь процесс вычисления $y(nT)$ реализуется с помощью трех базовых операций:

- задержки на один период дискретизации T ;
- умножения отсчета сигнала на постоянную величину;
- суммирования.

В элементе задержки сигнал на выходе запаздывает по отношению к входному на период T , но не изменяет других параметров и характеристик, например амплитуды, формы и т. д.

Временные характеристики цифровых фильтров. Основными характеристиками цифрового фильтра во временной области являются импульсная и переходная.

Импульсная характеристика (ИХ) $h(nT)$ представляет собой реакцию цифрового фильтра на цифровой единичный импульс $\delta(nT)$ при нулевых начальных условиях:

$$h(nT) = F[\delta(nT)].$$

ИХ может как иметь конечную длину, так и быть бесконечной. По этому признаку фильтры классифицируются на системы с конечной (КИХ-системы) и системы с бесконечной (БИХ-системы) импульсными характеристиками. Очевидно, что импульсная характеристика нерекурсивного фильтра имеет конечную длину, при этом отсчеты ИХ численно равны соответствующим коэффициентам ЛРУ нерекурсивного фильтра: $h(kT) = b_k$. Импульсная характеристика рекурсивного фильтра, как правило, имеет бесконечную длину. Однако в отдельных случаях рекурсивный фильтр может являться КИХ-системой.

Одним из удобств введения импульсной характеристики является возможность определения реакции ЦФ в виде аperiodической (линейной) свертки входной последовательности и ИХ:

$$y(nT) = \sum_{k=0}^n x(kT)h(nT - kT) = \sum_{k=0}^n x(nT - kT)h(kT). \quad (2.5)$$

Действительно, входной сигнал $x(nT)$ может быть представлен в виде $x(nT) = \sum_{k=0}^{\infty} x(kT)\delta(nT - kT)$. Так как реакцией на цифровой единичный импульс $\delta(nT)$ является импульсная характеристика $h(nT)$, то вследствие постоянства коэффициентов ЦФ (стационарности ЦФ) реакцией на последовательность $\delta(nT - kT)$ будет сигнал вида $h(nT - kT)$. Далее, на основе принципа суперпозиции, правомочного по причине линейности фильтра, можно утверждать, что реакцией на сигнал $\sum_{k=0}^{\infty} x(kT)\delta(nT - kT)$ будет последовательность

$$y(nT) = \sum_{k=0}^{\infty} x(kT)h(nT - kT) = \sum_{k=0}^{\infty} x(nT - kT)h(kT). \quad (2.6)$$

Правомерное предположение о том, что при $n < 0$ справедливы равенства $x(nT) = 0$ и $h(nT) = 0$, приводит выражение (2.6) к виду (2.5). Необходимо отметить, что формула (2.5) является аналогом определения реакции непрерывной системы в виде интеграла свертки: $y(t) = \int_0^t x(\tau)h(t - \tau) d\tau$.

Переходная характеристика (ПХ) $g(nT)$ ЦФ представляет собой реакцию фильтра на цифровой единичный скачок $u_1(nT)$ при нулевых начальных условиях:

$$g(nT) = F[u_1(nT)].$$

Справедливы следующие соотношения между импульсной и пе-

реходной характеристиками:

$$g(nT) = \sum_{k=0}^n h(kT); \quad (2.7)$$

$$h(nT) = g(nT) - g(nT - T). \quad (2.8)$$

Характеристика цифрового фильтра в z -области. Для представления цифрового фильтра в z -области необходимо применить z -преобразование к обеим частям уравнения (2.1). С учетом свойств линейности и сдвига z -преобразований полученное выражение имеет вид

$$\sum_{m=0}^{M-1} a_m Y(z) z^{-m} = \sum_{k=0}^{K-1} b_k X(z) z^{-k}.$$

Аналогично получаются соответствующие z -модели рекурсивного и нерекурсивного ЦФ:

$$Y(z) = - \sum_{m=1}^{M-1} a_m Y(z) z^{-m} + \sum_{k=0}^{K-1} b_k X(z) z^{-k}; \quad (2.9)$$

$$Y(z) = \sum_{k=0}^{K-1} b_k X(z) z^{-k}. \quad (2.10)$$

В z -области цифровой фильтр полностью характеризуется *передаточной функцией* (ПФ) $H(z)$, определяемой отношением z -преобразования выходного сигнала $y(nT)$ к z -преобразованию входного сигнала $x(nT)$ при нулевых начальных условиях

$$H(z) = \frac{Z\{y(nT)\}}{Z\{x(nT)\}} = \frac{Y(z)}{X(z)}. \quad (2.11)$$

Пример 2.1. На вход ЦФ поступает сигнал $x(nT) = \{1; -1, 3; 0,5; 0; 0; 0; \dots\}$. Реакция определяется как $y(nT) = \{3; 6; 0,5; 0; 0; 0; 0; \dots\}$. Необходимо определить передаточную функцию.

Z -изображения входной и выходной последовательностей имеют вид

$$X(z) = 1 - 1,3z^{-1} + 0,5z^{-2}; \quad Y(z) = 3 + 6z^{-1} + 0,5z^{-2}.$$

Следовательно, передаточная функция ЦФ выражается следующим образом:

$$H(z) = \frac{Y(z)}{X(z)} = \frac{3 + 6z^{-1} + 0,5z^{-2}}{1 - 1,3z^{-1} + 0,5z^{-2}}.$$

В данном примере $H(z)$ представляет собой ПФ РЦФ, однако данный ЦФ является КИХ-системой.

В соответствии с формулой (2.11) из уравнений (2.9) и (2.10) можно получить передаточные функции рекурсивного (*дробно-рациональная функция*) и нерекурсивного (*целая рациональная функ-*

ция) ЦФ соответственно:

$$H(z) = \frac{\sum_{k=0}^{K-1} b_k z^{-k}}{1 + \sum_{m=1}^{M-1} a_m z^{-m}}; \quad (2.12)$$

$$H(z) = \frac{b_0}{1 + \sum_{m=1}^{M-1} a_m z^{-m}}; \quad (2.13)$$

$$H(z) = \sum_{k=0}^{K-1} b_k z^{-k}. \quad (2.14)$$

Представляет интерес взаимосвязь ПФ и ИХ цифрового фильтра. Известно, что в случае поступления на вход ЦФ при нулевых начальных условиях цифрового единичного импульса справедливы равенства:

$$X(z) = Z[\delta(nT)] = 1; \quad Y(z) = Z[h(nT)].$$

Передаточная функция ЦФ может быть определена как z -преобразование от его импульсной характеристики:

$$H(z) = \frac{Y(z)}{X(z)} = Z[h(nT)] = \sum_{n=0}^{\infty} h(nT) z^{-n}.$$

Однозначность пары z -преобразований позволяет рассчитывать ИХ ЦФ по известной ПФ:

$$h(nT) = Z^{-1}[H(z)] = \frac{1}{2\pi j} \oint_C H(z) z^{n-1} dz.$$

Удобной формой представления передаточной функции служит определение ее особых точек: *нулей* и *полюсов*. Одно из следствий основной теоремы алгебры комплексных чисел утверждает, что многочлены $P(z)$ степени I , входящие в выражения (2.12)–(2.14), могут быть представлены произведением I линейных сомножителей, масштабированным постоянным коэффициентом B :

$$P(z) = B(z - z_1)(z - z_2)(z - z_3) \dots (z - z_i) \dots (z - z_I), \quad (2.15)$$

где z_i — вещественные или комплексные корни многочлена, обеспечивающие равенство $P(z) = 0$. Из формулы (2.15) видно, что общее число корней соответствует степени многочлена. Так, например, полином второй степени $P(z) = z^2 + a_1 z + a_2$ имеет, в зависимости от дискриминанта $D = a_1^2 - 4a_2$, либо два вещественных корня (при

$D > 0$):

$$z_{1,2} = \frac{-a_1 \pm \sqrt{D}}{2} = -\frac{a_1}{2} \pm \sqrt{\frac{a_1^2}{4} - a_2},$$

либо два комплексно-сопряженных корня (при $D < 0$):

$$z_{1,2} = \frac{-a_1 \pm \sqrt{D}}{2} = \sqrt{a_2} e^{\pm j \arccos\left(\frac{-a_1}{2\sqrt{a_2}}\right)}.$$

Корни многочлена числителя ПФ (уравнения $\sum_{k=0}^{K-1} b_k z^{-k} = 0$) называются *нулями* z_k^0 передаточной функции, а корни многочлена знаменателя (уравнения $1 + \sum_{m=1}^{M-1} a_m z^{-m} = 0$) — *полюсами* z_m^* передаточной функции. Очевидно, что ПФ ЦФ в точке нуля равна нулю, а в окрестностях точки полюса растет неограниченно. В целом расположение нулей и полюсов в плоскости z , зависящее от коэффициентов a_m и b_k , будет определять свойства (в том числе фильтрующие) ЦФ. Графически особые точки ПФ изображаются на комплексной карте нулей и полюсов, при этом полюс обычно изображается звездочкой, а ноль — кружочком.

Если коэффициенты a_m, b_k передаточной функции ЦФ являются вещественными, то особые точки ПФ могут быть вещественными или комплексно-сопряженными. Математический анализ, не рассматриваемый здесь, позволяет сделать ряд полезных выводов.

1. Если все полюсы передаточной функции ЦФ лежат внутри круга единичного радиуса на плоскости z , то импульсная характеристика $h(nT)$ ЦФ с течением времени затухает и стремится к нулю.

2. Если хотя бы один полюс передаточной функции ЦФ лежит вне круга единичного радиуса на плоскости z , то импульсная характеристика $h(nT)$ ЦФ с течением времени возрастает по абсолютной величине.

Пример 2.2. Определить нули и полюсы ЦФ с передаточной функцией

$$H(z) = \frac{-0,1 + 0,1z^{-1}}{1 + 0,5z^{-1}}.$$

Для определения нулей и полюсов ПФ ЦФ необходимо преобразовать ее к виду без отрицательных степеней:

$$H(z) = \frac{-0,1z + 0,1}{z + 0,5} = \frac{-0,1(z - 1)}{z + 0,5},$$

а нуль и полюс фильтра равны соответственно $z_1^0 = 1$, $z_1^* = -0,5$.

Карта нулей и полюсов исследуемого фильтра показана на рис. 2.5.

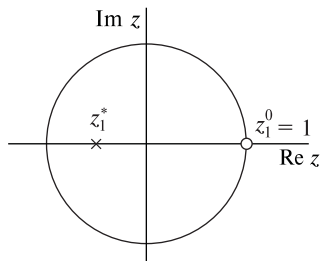


Рис. 2.5. Карта нулей и полюсов исследуемого фильтра

Устойчивость цифровых фильтров. Цифровой фильтр называется устойчивым, если при произвольных начальных условиях реакция $y(nT)$ на любое ограниченное воздействие $x(nT)$ также ограничена, т. е.:

если $|x(nT)| \leq M_x < \infty$ для всех $n = 0, 1, 2, \dots$,

то $|y(nT)| \leq M_y < \infty$ для всех $n = 0, 1, 2, \dots$

Тогда, с учетом формулы (2.5), справедливо соотношение

$$\begin{aligned} |y(nT)| &= \sum_{k=0}^{\infty} |h(kT)x(nT - kT)| \leq \sum_{k=0}^{\infty} |h(kT)||x(nT - kT)| \leq \\ &\leq M_x \sum_{k=0}^{\infty} |h(kT)|, \end{aligned}$$

на основе которого формулируется один из критериев устойчивости ЦФ: для устойчивости ЦФ необходима и достаточна абсолютная сходимость ряда отсчетов импульсной характеристики

$$\sum_{k=0}^{\infty} |h(kT)| < \infty. \quad (2.16)$$

Существует и другой, более простой для применения, критерий устойчивости ЦФ: все полюсы передаточной функции фильтра должны находиться внутри единичного круга z -плоскости, т. е.

$$|z_k^*| < 1. \quad (2.17)$$

Действительно, из формулы (2.16) следует, что

$$|H(z)| = \left| \sum_{k=0}^{\infty} h(kT)z^{-k} \right| \leq \sum_{k=0}^{\infty} |h(nT)||z^{-k}|.$$

При условии $|z| > 1$, что равносильно $|z^{-k}| < 1$, следует неравенство

$$|H(z)| \leq \sum_{k=0}^{\infty} |h(kT)|,$$

которое, с учетом формулы (2.16), для устойчивого ЦФ может быть дополнено до вида

$$|H(z)| \leq \sum_{k=0}^{\infty} |h(kT)| < \infty.$$

Следовательно, при указанных условиях ПФ конечна во всех точках z -плоскости, где $|z| < 1$. Это означает, что на единичной окружности z -плоскости и вне ее не должны находиться полюсы ПФ, в которых $H(z) = \infty$, т. е. должно выполняться условие (2.17). Необходимо отметить, что дробно-рациональная функция (2.12)–(2.13)

для анализа на устойчивость должна быть приведена к несокращаемому виду.

Расположение нулей ПФ не влияет на устойчивость ЦФ. Тогда, в соответствии с критерием (2.17), можно утверждать, что нерекурсивные ЦФ, являясь нулевыми, всегда устойчивы. Из критерия (2.16) следует, что всегда устойчивы КИХ-фильтры. Дополнительных исследований на устойчивость требуют те рекурсивные ЦФ, которые относятся к классу БИХ-фильтров, при этом условие устойчивости накладывает ограничения на значения коэффициентов полиномов, входящих в ПФ.

Пример 2.3. Оценить устойчивость ЦФ, конечная импульсная характеристика которого равна $h(nT) = \{10, -100, 8, -1, 0, 0, \dots\}$.

Для анализа используем первый критерий устойчивости. Так как

$$\sum_{n=0}^{\infty} |h(nT)| = 10 + 100 + 8 + 1 + 0 + 0 + \dots < \infty,$$

то этот КИХ-фильтр устойчив.

Пример 2.4. Определить устойчивость ЦФ с импульсной характеристикой $h(nT) = (-2)^n$.

Воспользуемся первым критерием устойчивости:

$$\sum_{n=0}^{\infty} |h(nT)| = 1 + 2 + 2^2 + 2^3 + \dots = \infty.$$

Полученный результат свидетельствует о неустойчивости исследуемого БИХ-фильтра.

Пример 2.5. ЦФ с передаточной функцией

$$H(z) = \frac{0,5}{1 - 0,9z^{-1}}$$

устойчив, так как в соответствии со вторым критерием устойчивости

$$|z_1^*| = 0,9 < 1.$$

Пример 2.6. Цифровой фильтр с передаточной функцией

$$H(z) = \frac{0,8}{1 - (0,9 - j0,95)z^{-1}}$$

неустойчив, так как в соответствии со вторым критерием устойчивости

$$|z_1^*| = \sqrt{(0,9)^2 + (0,95)^2} > 1.$$

Спектральные характеристики цифровых фильтров. В частотной области цифровой фильтр описывается частотной характеристикой (ЧХ) $H(e^{j\omega T})$, представляющей собой отношение преобразования Фурье (спектра) выходного сигнала к преобразованию Фурье (спектру) сигнала на входе фильтра в установившемся режиме функционирования:

$$H(e^{j\omega T}) = \frac{Y(e^{j\omega T})}{X(e^{j\omega T})}. \quad (2.18)$$

Кроме ЧХ, используются амплитудно-частотная, фазо-частотная характеристики и групповое время прохождения.

Частотная характеристика представляет собой частотную зависимость коэффициента передачи ЦФ и численно равна передаточной функции фильтра при $z = e^{j\omega T}$:

$$H(e^{j\omega T}) = H(z)|_{z=e^{j\omega T}}.$$

На основании формул (2.12) и (2.14) частотные характеристики рекурсивного и нерекурсивного фильтров соответственно имеют следующий вид:

$$H(e^{j\omega T}) = \frac{\sum_{k=0}^{K-1} b_k e^{-jk\omega T}}{1 + \sum_{m=1}^{M-1} a_m e^{-jm\omega T}}; \quad (2.19)$$

$$H(e^{j\omega T}) = \sum_{k=0}^{K-1} b_k e^{-jk\omega T}. \quad (2.20)$$

Частотная характеристика ЦФ в общем случае является комплексной функцией и поэтому ее можно представить в следующем виде:

$$H(e^{j\omega T}) = H_{\text{Re}}(e^{j\omega T}) + j H_{\text{Im}}(e^{j\omega T}) = A(\omega T) e^{j\varphi(\omega T)}.$$

Модуль частотной характеристики называется амплитудно-частотной характеристикой (АЧХ):

$$A(\omega T) = |H(e^{j\omega T})| = \sqrt{H_{\text{Re}}^2(e^{j\omega T}) + H_{\text{Im}}^2(e^{j\omega T})}, \quad (2.21)$$

а аргумент — фазочастотной характеристикой (ФЧХ) цифрового фильтра:

$$\varphi(\omega T) = \arg H(e^{j\omega T}) = \arctg \frac{H_{\text{Im}}(e^{j\omega T})}{H_{\text{Re}}(e^{j\omega T})}. \quad (2.22)$$

Общие формулы АЧХ и ФЧХ для РЦФ и НРЦФ получаются из выражений (2.19) и (2.20) на основании формул (2.21) и (2.22):

$$A_{\text{РЦФ}}(\omega T) = \sqrt{\frac{\left(\sum_{k=0}^{K-1} b_k \cos k\omega T\right)^2 + \left(\sum_{k=0}^{K-1} b_k \sin k\omega T\right)^2}{\left(1 + \sum_{m=1}^{M-1} a_m \cos m\omega T\right)^2 + \left(\sum_{m=1}^{M-1} a_m \sin m\omega T\right)^2}}; \quad (2.23)$$

$$A_{\text{НРЦФ}}(\omega T) = \sqrt{\left(\sum_{k=1}^{K-1} b_k \cos k\omega T\right)^2 + \left(\sum_{k=0}^{K-1} b_k \sin k\omega T\right)^2}; \quad (2.24)$$

$$\varphi_{\text{РЦФ}}(\omega T) = \operatorname{arctg} \left(\frac{-\sum_{k=0}^{K-1} b_k \sin k\omega T}{\sum_{k=0}^{K-1} b_k \cos k\omega T} \right) - \operatorname{arctg} \left(\frac{-\sum_{m=1}^{M-1} a_m \sin m\omega T}{1 + \sum_{m=1}^{M-1} a_m \cos m\omega T} \right); \quad (2.25)$$

$$\varphi_{\text{НРЦФ}}(\omega T) = \operatorname{arctg} \left(\frac{\sum_{k=0}^{K-1} b_k \sin k\omega T}{\sum_{k=0}^{K-1} b_k \cos k\omega T} \right). \quad (2.26)$$

Групповое время прохождения (ГВП) задается производной от ФЧХ, взятой с обратным знаком:

$$\tau(\omega T) = -\frac{d[\varphi(\omega T)]}{d(\omega T)}. \quad (2.27)$$

Ниже перечислены основные свойства частотных характеристик цифровых фильтров.

- Все частотные характеристики ЦФ являются непрерывными функциями частоты.
- Все частотные характеристики являются периодическими функциями с периодом, равным частоте дискретизации $\omega_d = 2\pi/T$.
- Если a_m и b_k — вещественные числа, то АЧХ и ГВЗ представляют собой четные функции частоты, а ФЧХ — нечетную функцию частоты.
- Если a_m и b_k — вещественные числа, то для полного описания частотной характеристики достаточно задать ее на интервале полупериода.

2.4. Классификация цифровых фильтров

Классификация цифровых фильтров производится по большому числу классификационных признаков. В частности, может быть применен набор признаков, использованный ранее при классификации дискретных систем в разд. 2.1. Еще один вариант классификации цифровых фильтров представлен табл. 2.1, некоторые пояснения к содержанию которой изложены ниже.

К линейным цифровым фильтрам относятся фильтры, описываемые линейным оператором F , т. е. такие, в которых выполняются принципы однородности и аддитивности, при этом входная $x(nT)$ и

Таблица 2.1

Вариант классификации цифровых фильтров

Классификационный признак		Классы цифровых фильтров
Тип оператора фильтрации		<ul style="list-style-type: none"> • линейные ЦФ • нелинейные ЦФ
Назначение		<ul style="list-style-type: none"> • частотно-избирательные (селективные) • специальные
Тип структуры	постоянство	<ul style="list-style-type: none"> • ЦФ с постоянной структурой • ЦФ с переменной структурой
	адаптивность	<ul style="list-style-type: none"> • неадаптивные ЦФ • структурно адаптивные ЦФ
Тип амплитудно-частотной характеристики селективного фильтра		<ul style="list-style-type: none"> • ЦФ нижних частот • ЦФ верхних частот • полосовые ЦФ • режекторные ЦФ
Тип фазочастотной характеристики		<ul style="list-style-type: none"> • ЦФ с линейной ФЧХ • ЦФ с нелинейной ФЧХ
Наличие обратных связей		<ul style="list-style-type: none"> • рекурсивные ЦФ • нерекурсивные ЦФ
Тип импульсной характеристики		<ul style="list-style-type: none"> • ЦФ с конечной ИХ (КИХ-фильтры) • ЦФ с бесконечной ИХ (БИХ-фильтры)
Размерность сигнального пространства		<ul style="list-style-type: none"> • одномерные ЦФ • многомерные ЦФ
Тип коэффициентов	постоянство	<ul style="list-style-type: none"> • ЦФ с постоянными коэффициентами • ЦФ с переменными коэффициентами
	вещественность	<ul style="list-style-type: none"> • ЦФ с вещественными коэффициентами • ЦФ с комплексными коэффициентами
	адаптивность	<ul style="list-style-type: none"> • параметрически адаптивные ЦФ • неадаптивные ЦФ

выходная $y(nT)$ последовательности фильтров связаны линейными разностными уравнениями. Фильтры, не отвечающие этому условию, относятся к классу нелинейных.

Наличие или отсутствие обратных связей классифицирует линейные ЦФ на рекурсивные (РЦФ) и нерекурсивные (НРЦФ) фильтры соответственно. Достаточно часто понятие нерекурсивности представляется термином «трансверсальность». Максимальная задержка входного сигнала, участвующего в формировании текущего отсчета $y(nT)$, выраженная числом периодов дискретизации $(K - 1)$, определяет порядок НРЦФ, в РЦФ порядок фильтра определяется аналогичным образом $(M - 1)$ по рекурсивной части фильтра.

Цифровые фильтры с постоянными параметрами характеризуются инвариантностью значений a_m и b_k во времени. В случае, когда хотя бы один из этих коэффициентов является функцией от пере-

менной n , ЦФ принадлежит к классу параметрических (фильтров с переменными параметрами). Достаточно часто в телекоммуникационных приложениях имеется необходимость изменять значения b_k и a_m по результатам анализа входной или выходной последовательностей. Такие параметрические ЦФ классифицируются как адаптивные с управлением по входу или по выходу соответственно.

К классу БИХ-фильтров относятся ЦФ, импульсная характеристика которых представляет собой бесконечную цифровую последовательность, а конечность ИХ классифицирует цифровые фильтры как КИХ-фильтры. Очевидно, что НРЦФ всегда относятся к фильтрам с конечной импульсной характеристикой, а рекурсивные ЦФ могут быть как БИХ- (в большинстве случаев), так и КИХ-фильтрами.

Следует отметить, что современное состояние теории и практики ЦОС делает возможным синтез и применение многомерных, нелинейных и адаптивных ЦФ, позволяющих реализовывать цифровые системы с характеристиками, недостижимыми на базе одномерных линейных цифровых фильтров с постоянными параметрами и постоянной структурой.

2.5. Структурные схемы фильтров

Структура цифрового фильтра отражает алгоритм обработки сигнала в фильтре и связана с видом ЛРУ и передаточной функции. Однако существуют различные структуры фильтров, обеспечивающие одну и ту же передаточную функцию. Основными структурами ЦФ являются *прямая, каноническая 1-го типа, каноническая 2-го типа и каскадная*. Под каноническими структурами подразумеваются такие, в которых используемое число элементов задержки в точности равно порядку передаточной функции (т. е. наивысшей степени многочленов числителя и знаменателя).

Прямая структура РЦФ второго порядка показана на рис. 2.3. Она соответствует ЛРУ и передаточной функции:

$$y(nT) = b_0x(nT) + b_1x(nT - T) + b_2x(nT - 2T) - a_1y(nT - T) - a_2y(nT - 2T); \quad (2.28)$$

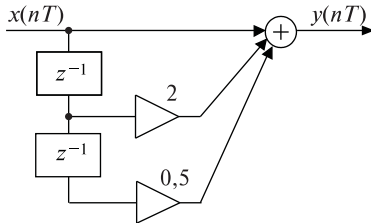
$$H(z) = \frac{b_0 + b_1z^{-1} + b_2z^{-2}}{1 + a_1z^{-1} + a_2z^{-2}}. \quad (2.29)$$

Аналогичные прямые структуры имеют фильтры меньшего (первого) и большего порядков. Особенностью РЦФ, имеющих прямую структуру, является наличие $K + M - 2$ однопериодных элементов задержки. Рисунок 2.3 является частным случаем построения прямой структуры при $K = M$, в более общем случае K может быть не равно M .

Структурная схема НРЦФ, соответствующая ЛРУ (2.3) и ПФ (2.17) представлена на рис. 2.4.

Пример 2.7. Построить структурную схему цифровой системы, если

$$H(z) = 1 + 2z^{-1} + 0,5z^{-2}.$$



В соответствии с определением передаточной функции можно записать:

$$H(z) = \frac{Y(z)}{X(z)} = 1 + 2z^{-1} + 0,5z^{-2},$$

тогда реакция системы в z -области имеет вид

$$Y(z) = X(z) + 2X(z)z^{-1} + 0,5X(z)z^{-2},$$

что соответствует линейному разностному уравнению

$$y(nT) = x(nT) + 2x(nT - 1T) + 0,5x(nT - 2T).$$

Рис. 2.6. Структурная схема НРЦФ второго порядка

Структурная схема исследуемой неrecursive системы приведена на рис. 2.6.

Каноническая структура 1-го типа РЦФ второго порядка. Для ее обоснования представим передаточную функцию (2.28) фильтра в виде

$$H(z) = \frac{1}{1 + a_1z^{-1} + a_2z^{-2}}(b_0 + b_1z^{-1} + b_2z^{-2}) = H_1(z)H_2(z),$$

что соответствует последовательному соединению РЦФ второго порядка с передаточной функцией

$$H_1(z) = \frac{Z[\nu(nT)]}{Z[x(nT)]} = \frac{1}{1 + a_1z^{-1} + a_2z^{-2}}$$

и НРЦФ второго порядка с передаточной функцией

$$H_2(z) = \frac{Z[y(nT)]}{Z[\nu(nT)]} = b_0 + b_1z^{-1} + b_2z^{-2}.$$

Соответствующие линейные разностные уравнения выражаются следующим образом:

$$\nu(nT) = x(nT) - a_1\nu(nT - T) - a_2\nu(nT - 2T);$$

$$y(nT) = b_0\nu(nT) + b_1\nu(nT - T) + b_2\nu(nT - 2T).$$

Тогда, используя РЦФ и НРЦФ второго порядка с прямыми структурами, можно представить структурную схему исследуемого РЦФ в виде, показанном на рис. 2.7. В точках 1 и 2 этой схемы сигналы одинаковы и равны $\nu(nT - T)$, а в точках 3 и 4 — одинаковы и равны $\nu(nT - 2T)$. Это дает основания упростить полученную схему.

Новая структурная схема, равнозначная схеме на рис. 2.7, изображена на рис. 2.8 и соответствует канонической структуре 1-го типа. Сравнение структур, представленных рис. 2.4 и 2.8, свидетельствует о том, что в случае канонической структуры 1-го типа

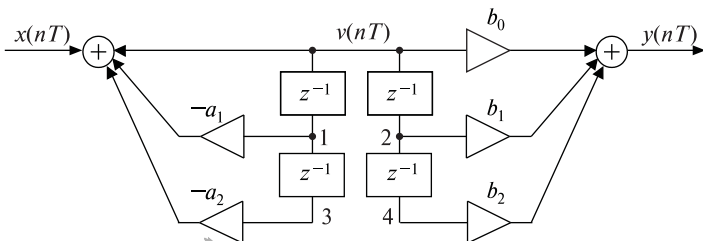


Рис. 2.7. Последовательное соединение РЦФ и НРЦФ второго порядка

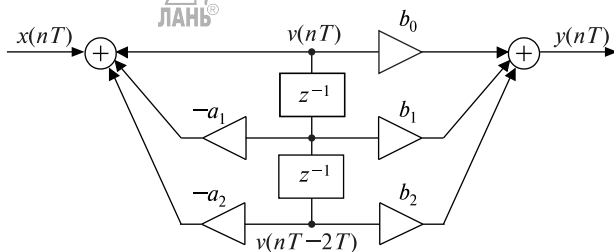


Рис. 2.8. Каноническая структура 1-го типа РЦФ второго порядка

увеличивается количество сумматоров, но уменьшается число элементов задержки.

Каноническая структура 2-го типа РЦФ второго порядка. Для ее обоснования преобразуем линейное разностное уравнение (2.33), введя в алгоритм обработки промежуточные дискретные сигналы $v(nT)$ и $d(nT)$:

$$y(nT) = b_0x(nT) + v(nT - T),$$

где

$$v(nT) = b_1x(nT) - a_1y(nT) + d(nT - T); \quad d(nT) = b_2x(nT) - a_2y(nT).$$

Уравнения описывают каноническую структуру 2-го типа для РЦФ второго порядка. Структурная схема показана на рис. 2.9.

Очевидно, что полученная структура вновь характеризуется минимально возможным количеством элементов задержки, но требует

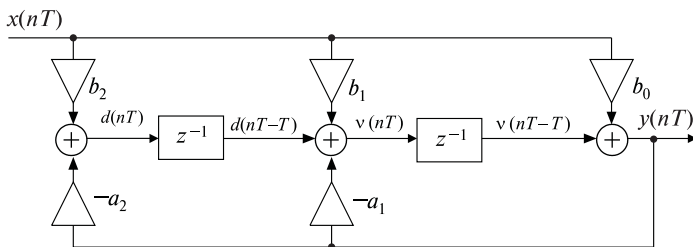


Рис. 2.9. Каноническая структура 2-го типа РЦФ второго порядка

большого, чем при канонической структуре 1-го типа, количества сумматоров.

Каскадная структура цифровых фильтров часто используется для построения фильтров больших порядков и представляет собой последовательное соединение фильтров низших порядков, реализованных в виде прямой или канонических структур. Расчет передаточной функции каскадных ЦФ рассмотрен ниже (пример 2.8). Примером каскадной структуры РЦФ второго порядка выступает структурная схема, изображенная на рис. 2.7.

Необходимо отметить, что цифровые фильтры, представленные на рис. 2.3, 2.7, 2.8 и 2.9, реализуют одну и ту же передаточную функцию (2.28).

В теории и практике цифровой фильтрации значительную роль играет понятие эквивалентности цифровых фильтров. *Эквивалентными* называются цифровые фильтры, реакция которых на одинаковые входные сигналы при нулевых начальных условиях одинакова. Одним из применений понятия эквивалентности является анализ передаточных функций $H_3(z)$ сложных цифровых фильтров, представляющих собой соединение более простых ЦФ с известными передаточными функциями $H_i(z)$. Возможно три типа соединений: последовательное, параллельное и соединение обратной связи.

Последовательное соединение фильтров с передаточными функциями $H_1(z)$ и $H_2(z)$ показано на рис. 2.10. Передаточная функция фильтра, эквивалентного представленному соединению, определяется следующим образом:

$$H_3(z) = \frac{X_3(z)}{X_1(z)}.$$

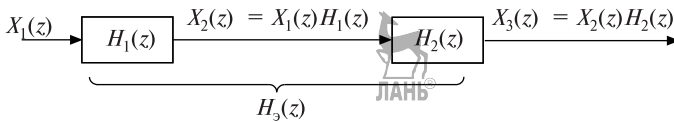


Рис. 2.10. Последовательное соединение цифровых фильтров

Z -изображение $X_3(z)$ выходного сигнала можно представить как $X_3(z) = X_1(z)H_1(z)H_2(z)$, тогда искомая передаточная функция принимает вид

$$H_3(z) = H_1(z)H_2(z).$$

В общем случае передаточная функция $H_3(z)$ фильтра, представляющего собой последовательное соединение I фильтров с частными передаточными функциями $H_i(z)$, $i = 1, 2, \dots, I$, равна произ-

ведению передаточных функций соединенных фильтров:

$$H_{\Sigma}(z) = \prod_{i=1}^I H_i(z).$$

Пример 2.8. Представить вариант синтеза фильтра с передаточной функцией

$$H_{\Sigma}(z) = \frac{b_0 + b_1 z^{-1}}{1 + a_1 z^{-1} + a_2 z^{-2}}$$

последовательным соединением более простых цифровых звеньев.

Требуемая передаточная функция может быть представлена

$$H_{\Sigma}(z) = \frac{1}{1 + a_1 z^{-1} + a_2 z^{-2}} (b_0 + b_1 z^{-1}) = H_1(z) H_2(z),$$

следовательно, фильтр с передаточной функцией $H_{\Sigma}(z)$ может быть сформирован последовательным соединением РЦФ с передаточной функцией $H_1(z) = \frac{1}{1 + a_1 z^{-1} + a_2 z^{-2}}$ и НРЦФ с передаточной функцией $H_2(z) = b_0 + b_1 z^{-1}$.

Параллельное соединение фильтров с передаточными функциями $H_1(z)$ и $H_2(z)$ показано на рис. 2.11. Передаточная функция фильтра, эквивалентного представленному соединению, определяется следующим образом:

$$H_{\Sigma}(z) = \frac{X_4(z)}{X_1(z)}.$$

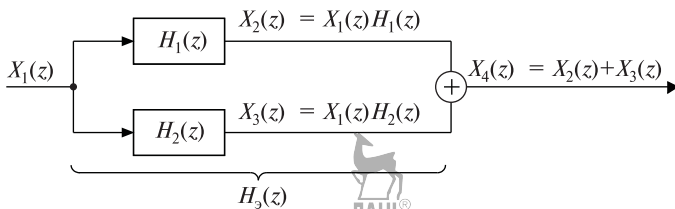


Рис. 2.11. Параллельное соединение цифровых фильтров

Z -изображение $X_4(z)$ выходного сигнала можно представить как $X_4(z) = X_2(z) + X_3(z) = X_1(z)[H_1(z) + H_2(z)]$, тогда искомая передаточная функция принимает вид $H_{\Sigma}(z) = H_1(z) + H_2(z)$.

В общем случае передаточная функция $H_{\Sigma}(z)$ фильтра, представляющего собой параллельное соединение I фильтров с частными передаточными функциями $H_i(z), i = 1, 2, \dots, I$, равна сумме передаточных функций соединенных фильтров:

$$H_{\Sigma}(z) = \sum_{i=1}^I H_i(z).$$

Соединение обратной связи фильтров с передаточными функциями $H_1(z)$ в прямой ветви и $H_2(z)$ в обратной ветви показано на рис. 2.12.

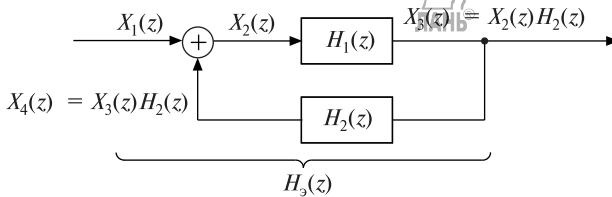


Рис. 2.12. Соединение обратной связи цифровых фильтров

Передаточная функция фильтра, эквивалентного представленному соединению, определяется следующим образом:

$$H_{\text{э}}(z) = \frac{X_3(z)}{X_1(z)}.$$

Z -изображение $X_3(z)$ выходного сигнала можно представить как

$$\begin{aligned} X_3(z) &= X_2(z)H_1(z) = [X_1(z) + X_3(z)H_2(z)]H_1(z) = \\ &= X_1(z)H_1(z) + X_3(z)H_1(z)H_2(z), \end{aligned}$$

а после преобразований как

$$X_3(z) = \frac{X_1(z)H_1(z)}{1 - H_1(z)H_2(z)},$$

тогда искомая передаточная функция принимает вид

$$H_{\text{э}}(z) = \frac{H_1(z)}{1 - H_1(z)H_2(z)}.$$

2.6. Простейшие цифровые фильтры и их характеристики

В теории ЦОС достаточно часто бывает удобным использовать при анализе и синтезе сложных цифровых систем простейшие цифровые фильтры. Ниже представлены примеры таких ЦФ: нерекурсивный и рекурсивные ЦФ первого и второго порядков. Следует иметь в виду, что в некоторых изданиях такие фильтры называются простейшими цифровыми звеньями.

НРЦФ 1-го порядка. Пусть линейно-разностное уравнение нерекурсивного цифрового фильтра 1-го порядка задано следующим выражением:

$$y(nT) = x(nT) - 0,8x(nT - T), \quad n = 0, 1, 2, \dots,$$

т.е. коэффициенты фильтра имеют следующие значения: $a_0 = 1$, $b_0 = 1$, $b_1 = -0,8$.

Прямая структура фильтра, описываемого данным ЛРУ, и сигналы, действующие в различных точках схемы, представлены на рис. 2.13.

Известно, что основными характеристиками цифровых фильтров во временной области являются импульсная и переходная характеристики.

Для расчета ИХ НРЦФ 1-го порядка можно воспользоваться его линейно-разностным уравнением с учетом следующих замен:

$$x(nT) = \delta(nT); \quad y(nT) = h(nT); \quad \delta(nT) = \begin{cases} 1, & n = 0; \\ 0, & n \neq 0. \end{cases}$$

Тогда общее выражение ИХ примет следующий вид:

$$h(nT) = \delta(nT) - 0,8\delta(nT - T).$$

Полученное выражение позволяет рассчитать отсчеты импульсной характеристики, соответствующие конкретным значениям времени:

$$\begin{aligned} h(0T) &= \delta(0T) - 0,8\delta(-T) = 1 - 0 = 1; \\ h(T) &= \delta(T) - 0,8\delta(0T) = 0 - 0,8 = -0,8; \\ h(2T) &= \delta(2T) - 0,8\delta(T) = 0. \end{aligned}$$

Очевидно, что при $n \geq 2$ отсчеты ИХ принимают нулевое значение. Графическое отображение рассчитанной ИХ представлено на рис. 2.14.

Таким образом, импульсная характеристика является конечной. Такой вывод обуславливает гарантированную устойчивость исследуемого цифрового фильтра.

Для расчета ПХ можно воспользоваться его линейно-разностным уравнением с учетом следующих замен:

$$\begin{cases} x(nT) = u_1(nT); \\ y(nT) = g(nT); \end{cases} \quad u_1(nT) = \begin{cases} 1, & n \geq 0; \\ 0, & n < 0. \end{cases}$$

Тогда общее выражение ПХ примет следующий вид:

$$g(nT) = u_1(nT) - 0,8u_1(nT - T).$$

Полученное выражение позволяет рассчитать отсчеты переходной характеристики, соответствующие конкретным значениям времени:

$$\begin{aligned} g(0T) &= u_1(0T) - 0,8u_1(-T) = 1 - 0 = 1; \\ g(T) &= u_1(T) - 0,8u_1(0T) = 1 - 0,8 = 0,2; \\ g(2T) &= u_1(2T) - 0,8u_1(T) = 1 - 0,8 = 0,2; \dots \end{aligned}$$

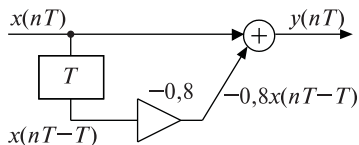


Рис. 2.13. Прямая структура нерекурсивного ЦФ 1-го порядка

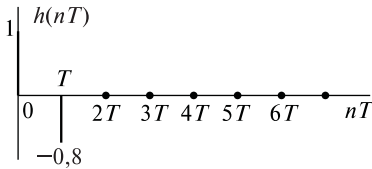


Рис. 2.14. Импульсная характеристика исследуемого НРЦФ

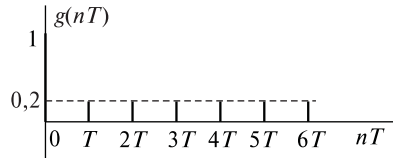


Рис. 2.15. Переходная характеристика исследуемого НРЦФ

Графическое отображение рассчитанной переходной характеристики представлено на рис. 2.15.

Полученные значения отсчетов ИХ и ПХ подтверждают справедливость выражений, определяющих взаимосвязь этих характеристик. Так, возможность определения переходной характеристики на основе импульсной можно продемонстрировать следующими расчетами:

$$g(0T) = h(0T) = 1;$$

$$g(T) = h(0T) + h(T) = 1 + (-0,8) = 0,2;$$

$$g(2T) = h(0T) + h(1T) + h(2T) = 1 + (-0,8) + 0 = 0,2;$$

$$g(3T) = h(0T) + h(1T) + h(2T) + h(3T) = 1 + (-0,8) + 0 + 0 = 0,2;$$

...

а обратная зависимость подтверждается так:

$$h(0T) = g(0T) = 1;$$

$$h(1T) = g(1T) - g(0T) = 0,2 - 1 = -0,8;$$

$$h(2T) = g(2T) - g(1T) = 0,2 - 0,2 = 0; \dots$$

Передаточная функция (ПФ) цифрового фильтра может быть получена применением прямого дискретного z -преобразования ко всем элементам ЛРУ (2.29). С учетом свойств дискретного z -преобразования это приводит к выражению $Y(z) = X(z) - 0,8X(z)z^{-1}$. Дальнейшее деление обеих частей полученного выражения на z -образ $X(z)$ входного сигнала позволяет получить передачную функцию:

$$\frac{Y(z)}{X(z)} = \frac{X(z) - 0,8X(z)z^{-1}}{X(z)} = 1 - 0,8z^{-1}; \quad H(z) = 1 - 0,8z^{-1}.$$

Полезным способом представления ПФ является карта нулей и полюсов. Для расчета нулей ПФ требуется выполнение условия $H(z) = 0$, что возможно лишь при $z = 0,8$. Таким образом, передачная функция имеет один нуль: $z_1^0 = 0,8$.

Расчет полюсов ПФ производится из условия $H(z) = \infty$, что для рассматриваемого ЦФ возможно лишь при $z_1^* = 0$, а это свойст-

венно всем нерекурсивным фильтрам. Таким образом, НРЦФ 1-го порядка, как и любой нерекурсивный цифровой фильтр, является гарантированно устойчивым.

Карта нулей и полюсов передаточной функции исследуемого ЦФ представлена на рис. 2.16.

Частотная характеристика (ЧХ) цифрового фильтра определяется выражением $H(e^{j\omega T}) = H(z)$ при $z = e^{j\omega T}$. Следовательно, ЧХ нерекурсивного фильтра 1-го порядка определяется следующим образом:

$$H(e^{j\omega T}) = 1 - 0,8e^{-j\omega T}. \quad (2.30)$$

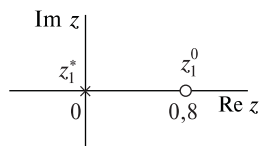


Рис. 2.16. Карта нулей и полюсов ПФ исследуемого фильтра

Комплексность полученного выражения затрудняет его анализ и использование в практических целях. На практике принято использовать модуль $|H(e^{j\omega T})|$ и аргумент $\arg H(e^{j\omega T})$ частотной характеристики, являющиеся вещественными функциями частоты и определяющие амплитудно-частотную (АЧХ) и фазочастотную (ФЧХ) характеристики цифрового фильтра соответственно. Периодичность АЧХ и ФЧХ позволяет проводить их расчет и анализ на длительности одного периода (например, на интервале от 0 до 1 по шкале нормированной частоты f_0). На практике достаточно часто исследования выполняются на длительности половины периода АЧХ и ФЧХ с учетом их четности и нечетности соответственно.

Следует иметь в виду, что практический интерес исследования состоит в анализе избирательных свойств фильтров на интервале частот от 0 до $\frac{f_d}{2}$ (от 0 до 0,5 по шкале нормированной частоты f_0), который называется «основной полосой частот». Понятие «основная полоса частот» возникло в результате того, что именно в этом частотном диапазоне лежит спектр исходного аналогового сигнала, подвергаемого дискретизации и ограниченного, согласно теореме В.А. Котельникова, верхней частотой $f_v \leq f_d/2$.

Для расчета АЧХ и ФЧХ удобно ввести следующие обозначения:

- $\text{Re}_H = 1 - 0,8 \cos \omega T$ — действительная часть выражения (2.30);
- $\text{Im}_H = 0,8 \sin \omega T$ — мнимая часть выражения (2.30).

АЧХ с учетом введенных обозначений определяется следующим образом:

$$\begin{aligned} A(\omega T) &= |H(e^{j\omega T})| = \sqrt{\text{Re}_H^2 + \text{Im}_H^2} = \\ &= \sqrt{(1 - 0,8 \cos \omega T)^2 + (0,8 \sin \omega T)^2}. \end{aligned}$$

Значения АЧХ являются вещественными коэффициентами передачи фильтра для каждого значения ωT в основной полосе частот

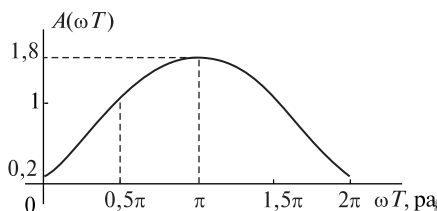


Рис. 2.17. Амплитудно-частотная характеристика исследуемого нерекурсивного ЦФ 1-го порядка

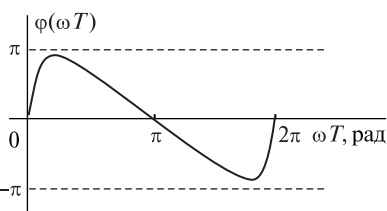


Рис. 2.18. Фазочастотная характеристика НРЦФ 1-го порядка

(по частотной шкале $\omega_0 = \omega T$ данная полоса частот определяется интервалом от 0 до π рад). Графическое представление рассчитанной АЧХ на длительности периода (на интервале от 0 до 2π по частотной шкале $\omega_0 = \omega T$) показано на рис. 2.17.

Конкретные значения АЧХ при известных значениях частоты могут быть получены расчетным путем. Ниже представлен расчет АЧХ для значений аргумента ωT , равных 0, $\pi/2$ и π радиан:

$$A(0) = \sqrt{[1 - 0,8 \cos 0]^2 + [-0,8 \sin 0]^2} = 0,2;$$

$$A\left(\frac{\pi}{2}\right) = \sqrt{\left[1 - 0,8 \cos \frac{\pi}{2}\right]^2 + \left[-0,8 \sin \frac{\pi}{2}\right]^2} \approx 1,28;$$

$$A(\pi) = \sqrt{[1 - 0,8 \cos \pi]^2 + [-0,8 \sin \pi]^2} = 1,8.$$

Фазочастотная характеристика НРЦФ 1-го порядка с учетом введенных обозначений определяется следующим образом:

$$\varphi(\omega T) = \arg H(e^{j\omega T}) = \arctg \frac{\operatorname{Im} H}{\operatorname{Re} H} = \arctg \frac{0,8 \sin \omega T}{1 - 0,8 \cos \omega T}.$$

Графическое представление рассчитанной ФЧХ показано на рис. 2.18.

Значения ФЧХ отражают изменения фаз частотных составляющих сигнала при его прохождении через фильтр. Очевидно, что ФЧХ исследуемого НРЦФ является нелинейной, следствием чего является неравномерность группового времени прохождения (ГВП) дискретных сигналов.

Конкретные значения ФЧХ при известных значениях частоты могут быть получены расчетным путем. Ниже представлены расчеты ФЧХ для некоторых значений аргумента ωT :

$$\varphi(0) = \arctg \frac{0,8 \sin 0}{1 - 0,8 \cos 0} = 0;$$

$$\varphi\left(\frac{\pi}{2}\right) = \arctg \frac{0,8 \sin \pi/2}{1 - 0,8 \cos \pi/2} = \arctg(0,8) \approx 0,67\pi;$$

$$\varphi(\pi) = \operatorname{arctg} \frac{0,8 \sin \pi}{1 - 0,8 \cos \pi} = 0.$$

Анализ характеристик исследуемого фильтра в спектральной области (АЧХ, ФЧХ, ГВП) свидетельствует о невыполнении условий безыскаженной передачи сигналов: АЧХ не является равномерной, ФЧХ не является линейной. Следовательно, дискретный сигнал при прохождении через исследуемый фильтр изменит свою форму. Очевидно, что изменение коэффициентов повлечет за собой изменение его характеристик.

РЦФ 1-го порядка. Приведенный ниже анализ рекурсивного фильтра 1-го порядка в основном аналогичен анализу НРЦФ 1-го порядка и по этой причине изложен более кратко. Пусть линейно-разностное уравнение рекурсивного цифрового фильтра 1-го порядка задано выражением

$$y(nT) = x(nT) - 0,8y(nT - T), \quad n = 0, 1, 2, \dots,$$

т.е. коэффициенты фильтра имеют значения: $a_0 = 1$, $a_1 = 0,8$, $b_0 = 1$.

Прямая структура звена, описываемого данным ЛРУ, и сигналы, действующие в различных точках схемы, представлены на рис. 2.19.

Общее выражение импульсной характеристики исследуемого фильтра имеет вид

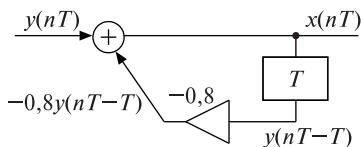


Рис. 2.19. Прямая структура исследуемого РЦФ

$$h(nT) = \delta(nT) - 0,8h(nT - T).$$

Полученное выражение позволяет рассчитать отсчеты ИХ, соответствующие конкретным значениям времени:

$$h(0T) = \delta(0T) - 0,8h(-T) = 1;$$

$$h(T) = \delta(T) - 0,8h(0T) = -0,8;$$

$$h(2T) = \delta(2T) - 0,8h(T) = (-0,8)^2 = 0,64;$$

$$h(3T) = \delta(3T) - 0,8h(2T) = (-0,8)^3 = -0,512;$$

.....

$$h(nT) = \delta(nT) - 0,8h(nT - T) = (-0,8)(-0,8)^{n-1} = (-0,8)^n.$$

Графическое отображение рассчитанной ИХ для $n = 0, 1, 2, 3$ представлено на рис. 2.20.

Анализ проведенных расчетов свидетельствует о том, что импульсная характеристика является бесконечной. Такой вывод обуславливает необходимость анализа исследуемого звена на устойчивость.

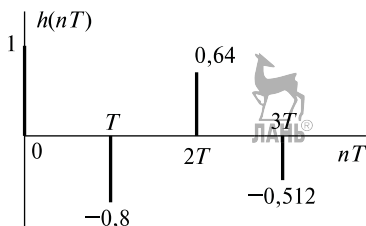


Рис. 2.20. Импульсная характеристика исследуемого РЦФ

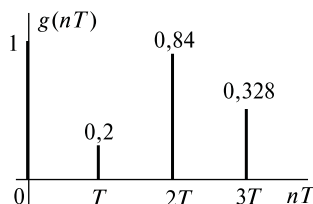


Рис. 2.21. Переходная характеристика исследуемого РЦФ

Общее выражение переходной характеристики имеет вид

$$g(nT) = u_1(nT) - 0,8g(nT - T).$$

Полученное выражение позволяет рассчитать отсчеты переходной характеристики, соответствующие конкретным значениям времени:

$$g(0T) = u_1(0T) - 0,8g(-T) = 1;$$

$$g(T) = u_1(T) - 0,8g(0T) = 1 - 0,8 = 0,2;$$

$$g(2T) = u_1(2T) - 0,8g(T) = 1 - 0,16 = 0,84;$$

$$g(3T) = u_1(3T) - 0,8g(2T) = 1 - 0,8 \cdot 0,84 = 1 - 0,672 = 0,328....$$

Графическое отображение рассчитанной переходной характеристики для $n = 0, 1, 2, 3$ представлено на рис. 2.21.

Известные зависимости между ИХ и ПХ можно продемонстрировать следующими расчетами:

$$g(0T) = h(0T) = 1;$$

$$g(T) = h(0T) + h(T) = 1 + (-0,8) = 0,2;$$

$$g(2T) = h(0T) + h(1T) + h(2T) = 1 + (-0,8) + 0,64 = 0,84;$$

$$\begin{aligned} g(3T) &= h(0T) + h(1T) + h(2T) + h(3T) = \\ &= 1 + (-0,8) + 0,64 + (-0,512) = 0,328; \end{aligned}$$

$$h(0T) = g(0T) = 1;$$

$$h(1T) = g(1T) - g(0T) = 0,2 - 1 = -0,8;$$

$$h(2T) = g(2T) - g(1T) = 0,84 - 0,2 = 0,64.$$

Передаточная функция (ПФ) исследуемого РЦФ может быть получена путем применения прямого дискретного z -преобразования ко всем элементам ЛРУ, что с учетом свойств дискретного z -преобразования приводит к выражению $Y(z) = X(z) - 0,8Y(z)z^{-1}$. Дальнейшее деление обеих частей полученного выражения на z -образ $X(z)$

входного сигнала позволяет получить передаточную функцию:

$$\begin{aligned}\frac{Y(z)}{X(z)} &= \frac{X(z) - 0,8Y(z)z^{-1}}{X(z)} = 1 - 0,8\frac{Y(z)}{X(z)}z^{-1}; \\ H(z) &= 1 - 0,8H(z)z^{-1}; \quad H(z)(1 + 0,8z^{-1}) = 1; \\ H(z) &= \frac{1}{1 + 0,8z^{-1}}.\end{aligned}$$

Для расчета координат особых точек z -плоскости (точек нулей и полюсов) целесообразно в выражении ПФ избавиться от отрицательной степени переменной z , представив его в следующем виде:

$$H(z) = \frac{1}{1 + 0,8z^{-1}} \frac{z}{z} = \frac{z}{z + 0,8}.$$

Очевидно, что дальнейший расчет нулей и полюсов ПФ связан с определением корней уравнений числителя и знаменателя передаточной функции.

Для расчета нулей требуется выполнение условия $H(z) = 0$, что возможно при $z = 0$. Таким образом, передаточная функция исследуемого фильтра имеет один нуль: $z_1^0 = 0$.

Расчет полюсов ПФ производится из условия $H(z) = \infty$, что возможно при $z + 0,8 = 0$. Следовательно, единственный полюс передаточной функции исследуемого рекурсивного фильтра 1-го порядка определяется значением $z_1^* = -0,8 = 0,8e^{j\pi}$.

Карта нулей и полюсов передаточной функции представлена на рис. 2.22.

Существенно, что точка полюса находится внутри единичной окружности z -плоскости, что в соответствии с критерием $|z_i^*| < 1$ подтверждает устойчивость исследуемого фильтра с бесконечной ИХ.

Частотная характеристика (ЧХ) определяется следующим образом:

$$H(e^{j\omega T}) = \frac{1}{1 + 0,8e^{-j\omega T}}.$$

Для вычисления АЧХ и ФЧХ исследуемого удобно ввести следующие обозначения:

- $\text{Re}_ч, \text{Im}_ч$ — действительная и мнимая части числителя выражения (2.30);
- $\text{Re}_з, \text{Im}_з$ — действительная и мнимая части знаменателя выражения (2.30).

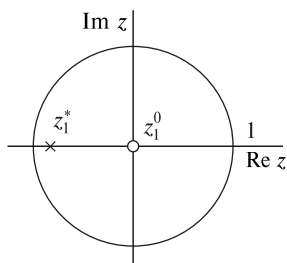


Рис. 2.22. Карта нулей и полюсов ПФ исследуемого фильтра

С учетом введенных обозначений АЧХ определяется следующим образом:

$$A(\omega T) = |H(e^{j\omega T})| = \frac{\sqrt{\operatorname{Re}_q^2 + \operatorname{Im}_q^2}}{\sqrt{\operatorname{Re}_3^2 + \operatorname{Im}_3^2}} = \frac{1}{\sqrt{[1 + 0,8 \cos \omega T]^2 + [0,8 \sin \omega T]^2}}.$$

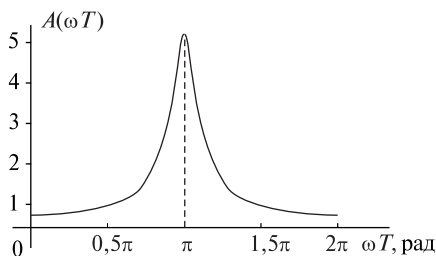


Рис. 2.23. Амплитудно-частотная характеристика исследуемого рекурсивного фильтра 1-го порядка

Графическое представление рассчитанной АЧХ на длительности периода (на интервале от 0 до 2π по частотной шкале $\omega_0 = \omega T$) показано на рис. 2.23.

Полученный график АЧХ свидетельствует о том, что в основной полосе частот РЦФ можно отнести к классу избирательных фильтров верхних частот (ФВЧ). Существенно, что это свойство присуще всем рекурсивным фильтрам, имеющим коэффициент $a_1 > 0$ (при $a_1 < 0$ такое звено является избирательным фильтром нижних частот). Плохое приближение полученной АЧХ к АЧХ идеального ФВЧ обусловлено первым порядком исследуемого ЦФ.

На практике при синтезе частотно-избирательных (селективных) ЦФ обеспечивается вычисление коэффициентов $\{a\}$ и $\{b\}$, гарантирующих непревышение АЧХ значения 1.

Конкретные значения АЧХ при известных значениях частоты могут быть получены расчетным путем. Ниже представлен расчет АЧХ для значений аргумента ωT , равных 0, $\pi/2$ и π радиан:

$$\begin{aligned} A(0) &= \sqrt{\frac{1}{[1 + 0,8 \cos 0]^2 + [0,8 \sin 0]^2}} \approx 0,56; \\ A\left(\frac{\pi}{2}\right) &= \sqrt{\frac{1}{[1 + 0,8 \cos \pi/2]^2 + [0,8 \sin \pi/2]^2}} \approx 0,78; \\ A(\pi) &= \sqrt{\frac{1}{[1 + 0,8 \cos \pi]^2 + [0,8 \sin \pi]^2}} \approx 5. \end{aligned}$$

Фазочастотная характеристика с учетом введенных обозначений определяется следующим образом:

$$\varphi(\omega T) = \arg H(e^{j\omega T}) = \operatorname{arctg} \frac{\operatorname{Im}_q}{\operatorname{Re}_q} - \operatorname{arctg} \frac{\operatorname{Im}_3}{\operatorname{Re}_3} =$$

$$= 0 - \operatorname{arctg} \frac{-0,8 \sin \omega T}{1 + 0,8 \cos \omega T} = \operatorname{arctg} \frac{0,8 \sin \omega T}{1 + 0,8 \cos \omega T}.$$

Графическое представление рассчитанной ФЧХ представлено на рис. 2.24. Очевидно, что ФЧХ является нелинейной, что определяет неравномерность группового времени прохождения (ГВП) дискретных сигналов через исследуемое звено.

Ниже представлены расчеты ФЧХ для некоторых значений аргумента ωT :

$$\varphi(0) = \operatorname{arctg} \frac{0,8 \sin 0}{1 + 0,8 \cos 0} = 0;$$

$$\varphi\left(\frac{\pi}{2}\right) = \operatorname{arctg} \frac{0,8 \sin \pi/2}{1 + 0,8 \cos \pi/2} \approx 0,67\pi;$$

$$\varphi\left(\frac{3\pi}{2}\right) = \operatorname{arctg} \frac{0,8 \sin 3\pi/4}{1 + 0,8 \cos 3\pi/4} \approx 0,916\pi;$$

$$\varphi(\pi) = \operatorname{arctg} \frac{0,8 \sin \pi}{1 + 0,8 \cos \pi} \approx 0.$$

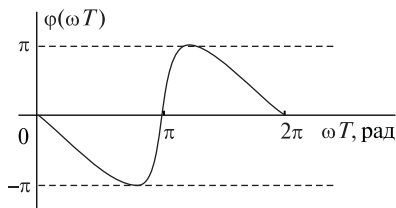


Рис. 2.24. Фазочастотная характеристика исследуемого рекурсивного фильтра 1-го порядка

Анализ характеристик исследуемого РЦФ в спектральной области (АЧХ, ФЧХ, ГВП) свидетельствует о невыполнении условий безыскаженной передачи сигналов: АЧХ не равномерная, ФЧХ нелинейная. Следовательно, дискретный сигнал при прохождении через такой фильтр изменит свою форму.

РЦФ 2-го порядка. Очевидным путем улучшения характеристик цифровых фильтров является повышение их порядка. Следует иметь в виду, что во многих изданиях цифровые фильтры 2-го порядков равноправно обозначаются терминами «базовое звено» и «базовый фильтр». По этой причине в дальнейшем материале для представления рекурсивного цифрового звена 2-го порядка будет использоваться термин «базовый рекурсивный фильтр 2-го порядка» (БРФ 2-го порядка).

Пусть линейно-разностное уравнение БРФ 2-го порядка задано следующим выражением:

$$y(nT) = x(nT) + 0,8y(nT - T) - 0,64y(nT - 2T), \quad n = 0, 1, 2, \dots$$

Прямая структура БРФ 2-го порядка, определяемого данным ЛРУ, и сигналы, действующие в различных точках схемы, представлены на рис. 2.25.

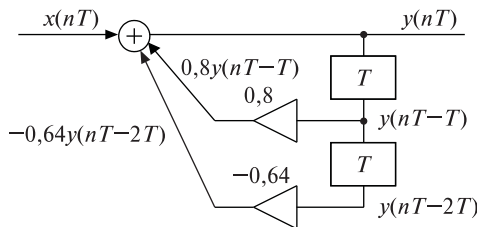


Рис. 2.25. Прямая структура исследуемого БРФ 2-го порядка

Импульсная характеристика исследуемого БРФ 2-го порядка определяется выражением

$$h(nT) = \delta(nT) + 0,8h(nT - T) - 0,64h(nT - 2T).$$

Расчет ее отсчетов приводит к следующим результатам:

$$h(0T) = \delta(0T) + 0,8h(-T) - 0,64h(0T - 2T) = 1;$$

$$h(T) = \delta(T) + 0,8h(0T) - 0,64h(T - 2T) = 0,8;$$

$$h(2T) = \delta(2T) + 0,8h(T) - 0,64h(0) = 0,64 - 0,64 = 0;$$

$$h(3T) = \delta(3T) + 0,8h(2T) - 0,64h(T) = 0,80 - 0,64 \cdot 0,8 = -0,512;$$

$$h(4T) = \delta(4T) + 0,8h(3T) - 0,64h(2T) = 0,8 \cdot (-0,512) = -0,4096; \dots$$

Анализ свидетельствует о том, что ИХ является бесконечной. Это делает обязательным анализ ЦФ на устойчивость.

Графическое представление ИХ исследуемого БРФ 2-го порядка для $n = 0, 1, 2, 3, 4$ показано на рис. 2.26.

Переходная характеристика исследуемого БРФ 2-го порядка определяется выражением

$$g(nT) = u_1(nT) + 0,8g(nT - T) - 0,64g(nT - 2T)$$

и является (как любая ПХ) бесконечной, а ее отсчеты для $n = 0, 1, 2, 3, 4$ имеют следующие значения:

$$g(0T) = u_1(0T) + 0,8g(-T) - 0,64g(-2T) = 1 + 0,8 \cdot 0 - 0,64 \cdot 0 = 1;$$

$$g(T) = u_1(T) + 0,8g(0T) - 0,64g(-T) = 1 + 0,8 \cdot 1 - 0,64 \cdot 0 = 1,8;$$

$$g(2T) = u_1(2T) + 0,8g(T) - 0,64g(0T) = 1 + 0,8 \cdot 1,8 - 0,64 \cdot 1 = 1,8;$$

$$g(3T) = u_1(3T) + 0,8g(2T) - 0,64g(T) = 1 + 0,8 \cdot 1,8 - 0,64 \cdot 1,8 = 1,28;$$

$$g(4T) = u_1(4T) + 0,8g(3T) - 0,64g(2T) = 1 + 0,8 \cdot 1,28 - 0,64 \cdot 1,8 = 0,87.$$

Графическое представление ПХ исследуемого БРФ 2-го порядка для $n = 0, 1, 2, 3, 4$ показано на рис. 2.27.

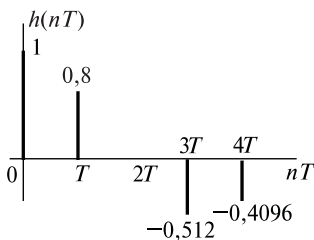


Рис. 2.26. Импульсная характеристика исследуемого БРФ 2-го порядка

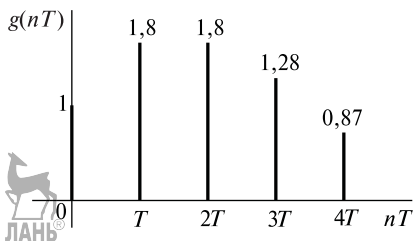


Рис. 2.27. Переходная характеристика исследуемого БРФ 2-го порядка

Передаточная функция исследуемого БРФ 2-го порядка может быть получена следующим образом:

$$\begin{aligned}
 Y(z) &= X(z) + 0,8Y(z)z^{-1} - 0,64Y(z)z^{-2}; \\
 H(z) &= \frac{Y(z)}{X(z)} = \frac{X(z) + 0,8Y(z)z^{-1} - 0,64Y(z)z^{-2}}{X(z)}; \\
 H(z) &= \frac{1}{1 - 0,8z^{-1} + 0,64z^{-2}}.
 \end{aligned}$$

Для определения нулей и полюсов ПФ удобно представить в следующем виде:

$$H(z) = \frac{z^2}{z^2 - 0,8z + 0,64}.$$

Анализ условия $H(z) = 0$ приводит к определению двух нулей ПФ: если $z^2 = 0$, то $z_{1,2}^0 = 0$.

Расчет полюсов ПФ предусматривает нахождение корней знаменателя ПФ, представленного полиномом второй степени.

Очевидно, что равенство данного полинома нулю: $z^2 - 0,8z + 0,64 = 0$ обеспечивает выполнение условия $H(z) = \infty$.

Дискриминант исследуемого полинома меньше нуля:

$$D = (-0,8)^2 - 4(0,64) = 0,64 - 2,56 = -1,92 < 0,$$

что обуславливает наличие двух комплексно-сопряженных корней, которые будут определять точки полюсов ПФ:

$$z_{1,2}^* = \frac{a_1 \pm \sqrt{D}}{a_2} = \sqrt{a_2} e^{\pm j \arccos\left(\frac{-a_1}{2\sqrt{a_2}}\right)},$$

где $a_1 = -0,8$, $a_2 = 0,64$. Тогда

$$\begin{aligned}
 z_{1,2}^* &= \sqrt{0,64} \exp \left[\pm j \arccos \left(\frac{0,8}{2\sqrt{0,64}} \right) \right] = 0,8 \exp [\pm j \arccos(0,5)] = \\
 &= 0,8 \exp [\pm j \pi / 3].
 \end{aligned}$$

Таким образом, карта нулей и полюсов ПФ исследуемого БРФ 2-го порядка имеет вид, показанный на рис. 2.28.

Полюса лежат внутри единичной окружности z -плоскости, что свидетельствует об устойчивости исследуемого БРФ 2-го порядка.

Частотная характеристика цифрового фильтра может быть получена по его передаточной функции при условии $z = e^{j\omega T}$. Следовательно, ЧХ исследуемого БРФ 2-го порядка определяется как

$$H(e^{j\omega T}) = \frac{1}{1 - 0,8e^{-j\omega T} + 0,64e^{-j2\omega T}}.$$

Неудобство практического использования ЧХ объясняется ее комплексным характером.

Амплитудно-частотная характеристика БРФ 2-го порядка определяется модулем его ЧХ:

$$\begin{aligned} A(\omega T) &= H|(e^{j\omega T})| = \sqrt{\frac{\operatorname{Re}_q^2 + \operatorname{Im}_q^2}{\operatorname{Re}_3^2 + \operatorname{Im}_3^2}} = \\ &= \sqrt{\frac{1}{[1 - 0,8 \cos \omega T + 0,64 \cos 2\omega T]^2 + [0,8 \sin \omega T - 0,64 \sin 2\omega T]^2}}. \end{aligned}$$

Графическое представление АЧХ в основной полосе частот показано на рис. 2.29.

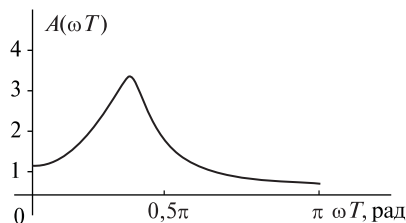


Рис. 2.29. АЧХ исследуемого БРФ 2-го порядка

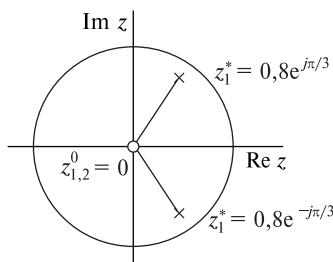


Рис. 2.28. Карта нулей и полюсов исследуемого БРФ 2-го порядка

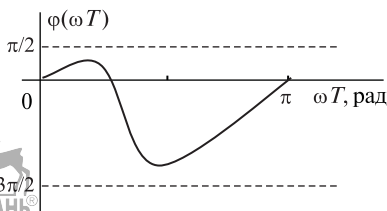


Рис. 2.30. Фазочастотная характеристика исследуемого БРФ 2-го порядка

Очевидно, что по типу избирательности исследуемый БРФ 2-го порядка может быть классифицирован как простейший фильтр нижних частот. Улучшение избирательных свойств фильтра достигается путем увеличения его порядка.

Фазочастотная характеристика исследуемого БРФ 2-го порядка определяется аргументом его частотной характеристики:

$$\begin{aligned}\varphi(\omega T) &= \arg H(e^{j\omega T}) = \\ &= \operatorname{arctg} \frac{\operatorname{Im}_ч}{\operatorname{Re}_ч} - \operatorname{arctg} \frac{\operatorname{Im}_з}{\operatorname{Re}_з} = - \operatorname{arctg} \frac{0,8 \sin \omega T - 0,64 \sin 2\omega T}{1 - 0,8 \cos \omega T + 0,64 \cos 2\omega T}.\end{aligned}$$

Графическое представление ФЧХ для основной полосы частот показано на рис. 2.30. Анализ ФЧХ свидетельствует о ее нелинейности, что обуславливает неравномерность группового времени прохождения (ГВП) дискретных сигналов через исследуемый фильтр.

Анализ характеристик исследуемого БРФ 2-го порядка в спектральной области (АЧХ, ФЧХ, ГВП) свидетельствует о невыполнении условий безыскаженной передачи сигналов: АЧХ не является равномерной, а ФЧХ — линейной. Следовательно, дискретный сигнал при прохождении через исследуемый фильтр изменит свою форму.

Представленные цифровые фильтры 1-го и 2-го порядков часто являются основой для каскадного построения частотно-избирательных и специальных фильтров более высоких порядков. Изменение коэффициентов базовых звеньев позволяет обеспечивать приближение реализуемых операторов фильтрации к требуемым.

2.7. Расчет характеристик рекурсивного цифрового фильтра во временной, частотной и z -областях

Ниже представлен расчет характеристик рекурсивного цифрового фильтра второго порядка, определяемого линейным разностным уравнением вида

$$y(n) = 0,6x(n) + 0,3x(n-1) - 0,4x(n-2) - 0,6y(n-1) - 0,3y(n-2).$$

Прямая структура такого фильтра показана на рис. 2.31. Очевидно, что коэффициенты a_m и b_k для данного фильтра имеют значения: $a_0 = 1$; $a_1 = 0,6$; $a_2 = 0,3$; $b_0 = 0,6$; $b_1 = 0,3$; $b_2 = -0,4$.

Расчет временных характеристик фильтра. Импульсная характеристика исследуемого фильтра

$$h(n) = 0,6\delta(n) + 0,3\delta(n-1) - 0,4\delta(n-2) - 0,6h(n-1) - 0,3h(n-2).$$

Отсчеты ИХ имеют значения:

$$h(0) = 0,6\delta(0) + 0,3\delta(-1) - 0,4\delta(-2) - 0,6h(-1) - 0,3h(-2) = 0,6;$$

$$\begin{aligned}h(1) &= 0,6\delta(1) + 0,3\delta(0) - 0,4\delta(-1) - 0,6h(0) - 0,3h(-1) = \\ &= 0,3 \cdot 1 - 0,6 \cdot 0,6 = -0,06;\end{aligned}$$

$$h(2) = 0,6\delta(2) + 0,3\delta(1) - 0,4\delta(0) - 0,6h(1) - 0,3h(0) =$$

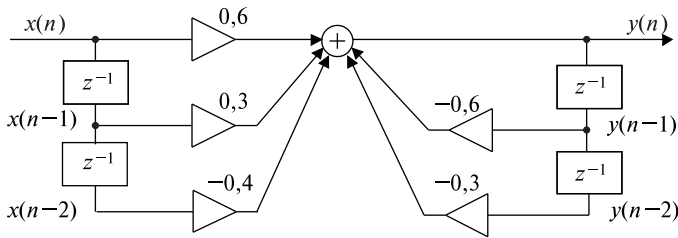


Рис. 2.31. Прямая структура исследуемого фильтра

$$= -0,4 \cdot 1 - 0,6 \cdot (-0,06) - 0,3 \cdot 0,6 = -0,544;$$

$$h(3) = 0,6\delta(3) + 0,3\delta(2) - 0,4\delta(1) - 0,6h(2) - 0,3h(1) = \\ = -0,6 \cdot (-0,544) - 0,3 \cdot (-0,06) = 0,3444;$$

Дальнейший анализ свидетельствует о бесконечности ИХ исследуемого фильтра. Такой вывод делает обязательным анализ устойчивости ЦФ. Расчет последующих отсчетов ИХ ($n > 3$) свидетельствует о соблюдении признака $\sum_{n=0}^{\infty} |h(nT)| < \infty$, следовательно, исследуемый РЦФ является устойчивым БИХ-фильтром. Графическое представление рассчитанной ИХ показано на рис. 2.32.

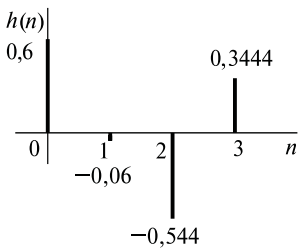


Рис. 2.32. Импульсная характеристика исследуемого ЦФ

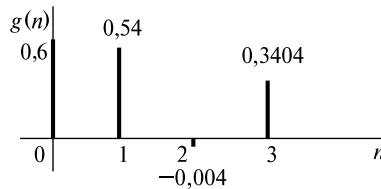


Рис. 2.33. Переходная характеристика исследуемого ЦФ

Переходная характеристика ЦФ

$$g(n) = 0,6u_1(n) + 0,3u_1(n-1) - 0,4u_1(n-2) - 0,6g(n-1) - 0,3g(n-2),$$

следовательно, отсчеты ПХ имеют значения:

$$g(0) = 0,6u_1(0) + 0,3u_1(-1) - 0,4u_1(-2) - 0,6g(-1) - 0,3g(-2) = 0,6;$$

$$g(1) = 0,6u_1(1) + 0,3u_1(0) - 0,4u_1(-1) - 0,6g(0) - 0,3g(-1) = \\ = 0,6 \cdot 1 + 0,3 \cdot 1 - 0,6 \cdot 0,6 = 0,54;$$

$$g(2) = 0,6u_1(2) + 0,3u_1(1) - 0,4u_1(0) - 0,6g(1) - 0,3g(0) = \\ = 0,6 \cdot 1 + 0,3 \cdot 1 - 0,4 \cdot 1 - 0,6 \cdot 0,54 - 0,3 \cdot 0,6 = -0,004;$$

$$g(3) = 0,6u_1(3) + 0,3u_1(2) - 0,4u_1(1) - 0,6g(2) - 0,3g(1) = \\ = 0,6 \cdot 1 + 0,3 \cdot 1 - 0,4 \cdot 1 - 0,6 \cdot (-0,004) - 0,3 \cdot 0,54 = 0,3404.$$

График переходной характеристики представлен на рис. 2.33.

Полученные значения отсчетов импульсной и переходной характеристик подтверждают справедливость выражений (2.7) и (2.8). Так, например:

$$\begin{aligned} g(0) &= h(0) = 0,6; \\ g(1) &= h(0) + h(1) = 0,6 - 0,06 = 0,54; \\ g(2) &= h(0) + h(1) + h(2) = 0,6 - 0,06 - 0,544 = -0,004; \\ h(1) &= g(1) - g(0) = 0,54 - 0,6 = -0,06; \\ h(2) &= g(2) - g(1) = -0,004 - 0,54 = -0,544 \text{ и т. д.} \end{aligned}$$

Расчет передаточной функции фильтра. Одним из путей для определения ПФ является выполнение прямого z -преобразования над линейным разностным уравнением, описывающим исследуемый ЦФ:

$$Y(z) = 0,6X(z) + 0,3X(z)z^{-1} - 0,4X(z)z^{-2} - 0,6Y(z)z^{-1} - 0,3Y(z)z^{-2}.$$

Незначительные преобразования полученного выражения позволяют получить искомую зависимость:

$$\begin{aligned} Y(z) (1 + 0,6z^{-1} + 0,3z^{-2}) &= X(z) (0,6 + 0,3z^{-1} - 0,4z^{-2}); \\ H(z) = \frac{Y(z)}{X(z)} &= \frac{0,6 + 0,3z^{-1} - 0,4z^{-2}}{1 + 0,6z^{-1} + 0,3z^{-2}}. \end{aligned}$$

Одним из полезных способов представления ПФ является карта ее нулей и полюсов. Для расчета нулей и полюсов передаточной функции необходимо, как правило, избавиться от отрицательных степеней переменной z :

$$H(z) = \frac{0,6 + 0,3z^{-1} - 0,4z^{-2}}{1 + 0,6z^{-1} + 0,3z^{-2}} \cdot \frac{z^2}{z^2} = \frac{0,6z^2 + 0,3z - 0,4}{z^2 + 0,6z + 0,3}$$

и определить корни полиномов числителя и знаменателя получившейся дробно-рациональной функции.

Расчет полюсов передаточной функции требует решения квадратного уравнения вида $z^2 + 0,6z + 0,3 = 0$. Дискриминант данного уравнения $D = 0,6^2 - 4 \cdot 0,3 = -0,84$ меньше нуля, что свидетельствует о наличии у ПФ двух комплексно-сопряженных полюсов:

$$\begin{aligned} z_1^* &= \frac{-0,6 + j\sqrt{0,84}}{2} \approx -0,3 + j0,4582 \approx 0,548e^{j2,15}; \\ z_2^* &= \frac{-0,6 - j\sqrt{0,84}}{2} \approx -0,3 - j0,4582 \approx 0,548e^{-j2,15}. \end{aligned}$$

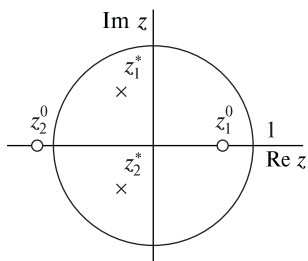


Рис. 2.34. Карта нулей и полюсов ПФ исследуемого фильтра

Расчет нулей передаточной функции требует решения квадратного уравнения вида $0,6z^2 + 0,3z - 0,4 = 0$ или, что равнозначно, $z^2 + 0,5z - 2/3 = 0$. Дискриминант данного уравнения $D = 0,5^2 + 8/3 \approx 2,92$ больше нуля, что свидетельствует о наличии у ПФ двух вещественных нулей:

$$z_1^0 = -\frac{0,5}{2} + \sqrt{\frac{0,25}{4} + \frac{2}{3}} \approx 0,6 \approx 0,6e^{j0};$$

$$z_2^0 = -\frac{0,5}{2} - \sqrt{\frac{0,25}{4} + \frac{2}{3}} \approx -1,1 \approx 1,1e^{j\pi}.$$

Карта нулей и полюсов фильтра показана на рис. 2.34.

Оба полюса передаточной функции лежат внутри единичной окружности z -плоскости, что, в соответствии с критерием (2.17), подтверждает устойчивость исследуемого рекурсивного цифрового фильтра с бесконечной импульсной характеристикой.

Расчет спектральных характеристик фильтра. Частотная характеристика ПФ определяется выражением (2.18) и для исследуемого рекурсивного фильтра соответствует формуле (2.19):

$$H(e^{j\omega T}) = \frac{0,6 + 0,3e^{-j\omega T} - 0,4e^{-j2\omega T}}{1 + 0,6e^{-j\omega T} + 0,3e^{-j2\omega T}}.$$

Амплитудно-частотная характеристика исследуемого фильтра определяется на основе выражений (2.21) и (2.22):

$$\begin{aligned} A(\omega T) &= |H(e^{j\omega T})| = \\ &= \left| \frac{0,6 + 0,3 \cos \omega T - 0,3j \sin \omega T - 0,4 \cos 2\omega T + 0,4j \sin 2\omega T}{1 + 0,6 \cos \omega T - 0,6j \sin \omega T + 0,3 \cos 2\omega T - 0,3j \sin 2\omega T} \right| = \\ &= \sqrt{\frac{(0,6 + 0,3 \cos \omega T - 0,4 \cos 2\omega T)^2 + (-0,3 \sin \omega T + 0,4 \sin 2\omega T)^2}{(1 + 0,6 \cos \omega T + 0,3 \cos 2\omega T)^2 + (-0,6 \sin \omega T - 0,3 \sin 2\omega T)^2}}. \end{aligned}$$

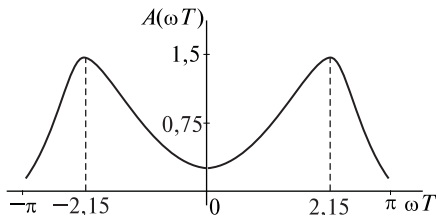


Рис. 2.35. Амплитудно-частотная характеристика исследуемого фильтра

Один период полученной зависимости графически показан на рис. 2.35. Его анализ свидетельствует о неравномерности масштабирования фильтром спектральных составляющих сигнала в основной полосе частот.

Соотношение между аргументами передаточной функции и частотной характеристики цифрового фильтра ($z = e^{j\varphi} = e^{j\omega T}$) делает возможным расчет значений

АЧХ на частотах, соответствующих фазовым углам нулей и полюсов передаточной функции:

$$\varphi_1^* \approx 2,15 \text{ рад}; \quad \varphi_2^* \approx -2,15 \text{ рад}; \quad \varphi_1^0 = 0 \text{ рад}; \quad \varphi_2^0 = \pi \text{ рад};$$

$$A_1^*(\omega T) \approx$$

$$\approx \sqrt{\frac{(0,6 + 0,3 \cos 2,15 - 0,4 \cos 4,3)^2 + (-0,3 \sin 2,15 + 0,4 \sin 4,3)^2}{(1 + 0,6 \cos 2,15 + 0,3 \cos 4,3)^2 + (-0,6 \sin 2,15 - 0,3 \sin 4,3)^2}} \approx$$

$$\approx 1,439; \quad A_2^*(\omega T) \approx$$

$$\approx \sqrt{\frac{(0,6 + 0,3 \cos 2,15 - 0,4 \cos 4,3)^2 + (0,3 \sin 2,15 - 0,4 \sin 4,3)^2}{(1 + 0,6 \cos 2,15 + 0,3 \cos 4,3)^2 + (0,6 \sin 2,15 + 0,3 \sin 4,3)^2}} \approx$$

$$\approx 1,439; \quad A_1^0(\omega T) \approx$$

$$\approx \sqrt{\frac{(0,6 + 0,3 \cos 0 - 0,4 \cos 0)^2 + (-0,3 \sin 0 + 0,4 \sin 0)^2}{(1 + 0,6 \cos 0 + 0,3 \cos 0)^2 + (-0,6 \sin 0 - 0,3 \sin 0)^2}} \approx$$

$$\approx 0,263; \quad A_2^0(\omega T) \approx$$

$$\approx \sqrt{\frac{(0,6 + 0,3 \cos \pi - 0,4 \cos 2\pi)^2 + (-0,3 \sin \pi + 0,4 \sin 2\pi)^2}{(1 + 0,6 \cos \pi + 0,3 \cos 2\pi)^2 + (-0,6 \sin \pi - 0,3 \sin 2\pi)^2}} \approx 0,143.$$

Анализ показывает, что полученные значения $A_1^*(\omega T)$, $A_2^*(\omega T)$, $A_1^0(\omega T)$ и $A_2^0(\omega T)$ достаточно близки к соответствующим точкам максимумов и минимумов АЧХ исследуемого фильтра.

Фазочастотная характеристика исследуемого фильтра определяется на основе выражений (2.22) и (2.25):

$$\varphi(\omega T) = \operatorname{arctg} \frac{-0,3 \sin \omega T + 0,4 \sin 2\omega T}{0,6 + 0,3 \cos \omega T - 0,4 \cos 2\omega T} -$$

$$- \operatorname{arctg} \frac{-0,6 \sin \omega T - 0,3 \sin 2\omega T}{1 + 0,6 \cos \omega T + 0,3 \cos 2\omega T}$$

и графически представлена на рис. 2.36.

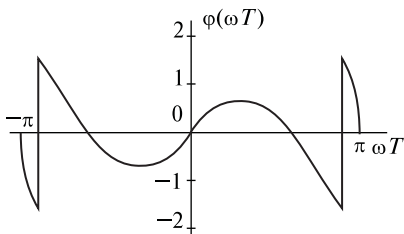


Рис. 2.36. Фазочастотная характеристика исследуемого фильтра

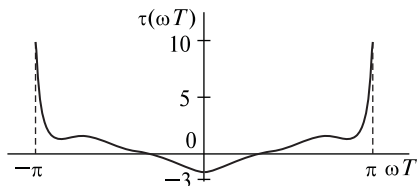


Рис. 2.37. Групповое время прохождения исследуемого фильтра

Полученная зависимость свидетельствует о принадлежности исследуемого фильтра к классу фильтров с нелинейной ФЧХ.

Групповое время прохождения фильтра определяется на основе формулы (2.27) и графически представлено на рис. 2.37.

Полученные результаты расчетов позволяют классифицировать исследуемый рекурсивный цифровой фильтр как устойчивый полосовой БИХ-фильтр с нелинейной ФЧХ.

2.8. Проектирование цифровых фильтров с конечной импульсной характеристикой

Особенности проектирования и применения фильтров связаны не с наличием или отсутствием обратной связи (рекурсивностью или нерекурсивностью ЦФ), а с видом импульсной характеристики (ее конечностью или бесконечностью). По этой причине для определения класса фильтров при решении задач проектирования чаще используются термины «БИХ-фильтр» и «КИХ-фильтр».

Ниже перечислены основные этапы проектирования ЦФ с конечной импульсной характеристикой.

Этап 1. Формулировка задачи аппроксимации.

Целью этапа является формулировка математической задачи, связывающей коэффициенты b_k КИХ-фильтра с требованиями к его характеристикам.

1.1. Определение аппроксимируемой функции, задающей требо-

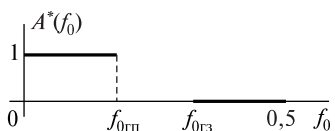


Рис. 2.38. Пример требований к АЧХ цифрового ФНЧ

вания к избранной характеристике ЦФ. На рис. 2.38 показан пример задания функции $A^*(f_0)$, определяющей требования к АЧХ цифрового ФНЧ:

$$A^*(f_0) = \begin{cases} 1, & 0 \leq f_0 \leq f_{0гп}; \\ 0, & f_{0гз} \leq f_0 \leq 0,5 f_0. \end{cases}$$

Между граничными частотами полосы пропускания $f_{0гп}$ и полосы задерживания $f_{0гз}$ располагается полоса расфильтровки, в которой также могут задаваться требования к АЧХ фильтра.

1.2. Выбор типа фильтра. Особый интерес представляют минимально-фазовые КИХ-фильтры и КИХ-фильтры с линейной ФЧХ.

Популярность минимально-фазовых КИХ-фильтров обусловлена минимальностью группового времени замедления, обеспечиваемой такими фильтрами. Особенностью передаточных функций минимально-фазовых фильтров является размещение их нулей внутри и на единичной окружности z -плоскости.

Основным достоинством КИХ-фильтров с линейной ФЧХ является сохранение формы сигнала, спектр которого полностью укладывается в полосу пропускания фильтра, где его коэффициент передачи постоянен. При проектировании КИХ-фильтров с линейной ФЧХ используются фильтры с симметричными и антисимметричными импульсными характеристиками относительно их середины. Они хорошо изучены, некоторые свойства таких ЦФ представлены в табл. 2.2.

Таблица 2.2

Некоторые свойства КИХ-фильтров с линейной ФЧХ

Свойство	Тип фильтра			
	1	2	3	4
Количество отсчетов конечной ИХ	Нечетное	Четное	Нечетное	Четное
Тип симметрии конечной ИХ	Симметричная	Симметричная	Антисимметричная	Антисимметричная
Определение $h(n)$, $n = 0, 1, 2, \dots, K - 1$	$h(K - 1 - n)$	$h(K - 1 - n)$	$-h(K - 1 - n)$	$-h(K - 1 - n)$
Вид ИХ	$K = 5$	$K = 6$	$K = 5$	$K = 6$
Возможные классы фильтров по типу АЧХ	ФНЧ, ФВЧ, ПФ, РФ	ФНЧ, ПФ	ПФ, дифференциаторы, ЦПФ	ФВЧ, ПФ, дифференциаторы, ЦПФ

1.3. Выбор аппроксимирующей функции $R(f_0, \mathbf{c})$, где \mathbf{c} — вектор коэффициентов, определяемый совокупностью коэффициентов b_k , а в более общем случае — связанный с ними.

1.4. Выбор весовой функции $q(f_0)$, управляющей точностью аппроксимации в различных частотных полосах. Весовая функция

$$q(f_0) = \begin{cases} 1, & 0 \leq f_0 \leq f_{0\text{гп}}, f_{0\text{гз}} \leq f_0 \leq 0,5; \\ 0, & f_{0\text{гп}} < f_0 < f_{0\text{гз}}, \end{cases}$$

детерминирует равнозначность ошибок аппроксимации АЧХ в полосах пропускания и задерживания и отсутствие влияния ошибок аппроксимации в полосе расфильтровки.

1.5. Выбор критерия аппроксимации, уточняющего смысл приближенного равенства $R(f_0, \mathbf{c}) \approx A^*(f_0)$ при $f_0 \in [f_{01}, f_{02}]$. Очевидно, что для большинства задач $f_{01} = 0$, $f_{02} = 0,5$. Наиболее часто используются среднеквадратический критерий

$$\int_{f_{01}}^{f_{02}} q(f_0) [A^*(f_0) - R(f_0, \mathbf{c})]^2 df_0 \rightarrow \min$$

или равномерный (чебышевский) критерий, минимизирующий модуль максимальной взвешенной ошибки аппроксимации $\varepsilon(f_0, \mathbf{c})$:

$$\max_{f_0 \in [f_{01}, f_{02}]} [\varepsilon(f_0, \mathbf{c})] = \max_{f_0 \in [f_{01}, f_{02}]} \{q(f_0) |A^*(f_0) - R(f_0, \mathbf{c})|\} \rightarrow \min. \quad (2.31)$$

Функция $R(f_0, \mathbf{c})$, удовлетворяющая условию (2.31), называется функцией наилучшего равномерного приближения.

Этап 2. Решение задачи аппроксимации.

Целью этапа является вычисление коэффициентов b_k , обеспечивающих требуемые характеристики КИХ-фильтра.

2.1. Определение необходимого (как правило, минимально допустимого) порядка $(K-1)$ КИХ-фильтра. Основные способы — использование эмпирических формул, содержащихся в руководствах по ЦОС, или подбор методом многократного повторения процедур проектирования фильтра.

2.2. Расчет вектора коэффициентов \mathbf{c} .

2.3. Проверка выполнения требований к характеристикам проектируемого КИХ-фильтра.

2.4. В случае, если полученный вектор \mathbf{c} обеспечивает выполнение требований к характеристикам проектируемого КИХ-фильтра — расчет коэффициентов b_k , в противном случае — возвращение к пункту 2.1 для увеличения значения K .

Этап 3. Расчет разрядности представления сигналов, коэффициентов, требуемого объема памяти.

Этап 4. Выбор элементной базы, разработка вычислителя.

Этап 5. Разработка программной реализации КИХ-фильтра.

Этап 6. Оценка характеристик реализованного КИХ-фильтра.

Синтез цифровых КИХ-фильтров с линейной фазочастотной характеристикой. Основными методами проектирования КИХ-фильтров с линейной ФЧХ являются метод взвешивания с помощью оконных функций, метод наименьших квадратов, метод частотной выборки, метод чебышевской аппроксимации и метод быстрой свертки. Ниже рассмотрены постановка и решение методом чебышевской аппроксимации задачи синтеза ФНЧ на основе нерекурсивного ЦФ с линейной ФЧХ, относящегося к типу 1 (см. табл. 2.2).

Частотная характеристика такого фильтра имеет вид

$$H(e^{j2\pi f_0}) = \sum_{k=0}^{K-1} b_k e^{-jk2\pi f_0} = e^{-j\pi f_0(K-1)} R(f_0, \mathbf{c}), \quad (2.32)$$

где

$$R(f_0, \mathbf{c}) = \sum_{k=0}^L c_k \cos k2\pi f_0; \quad (2.33)$$

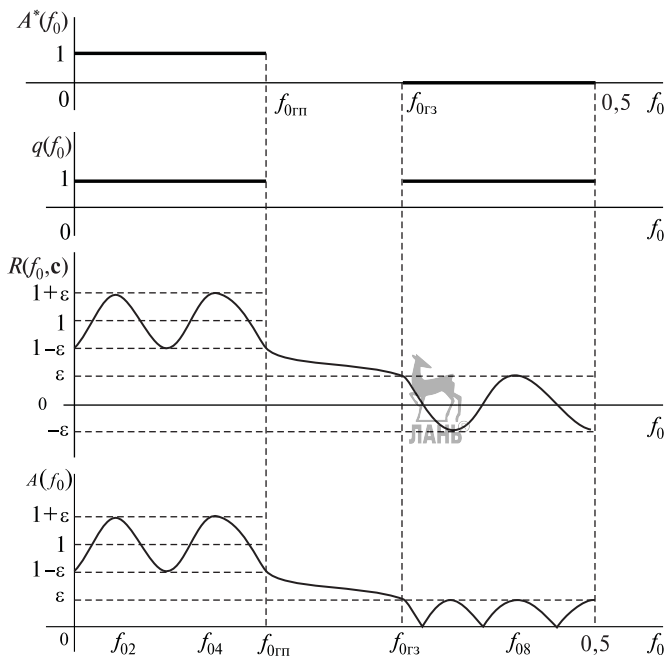


Рис. 2.39. Графическая иллюстрация решения задачи чебышевской аппроксимации АЧХ цифрового ФНЧ на основе КИХ-фильтра

$$L = \frac{K-1}{2}; \quad \mathbf{c} = (c_0, c_1, \dots, c_k, \dots, c_L); \quad c_k = \begin{cases} b_L, & k = 0; \\ 2b_{L-k}, & k = 1, 2, \dots, L. \end{cases}$$

В соответствии с формулой (2.32) амплитудно-частотная характеристика синтезируемого ЦФ определяется выражением $A(f_0) = |R(f_0, \mathbf{c})|$.

Решение аппроксимационной задачи заключается в определении коэффициентов b_k , обеспечивающих необходимую степень близости реальных характеристик ЦФ к заданным. Очевидно, что промежуточным этапом решения задачи должно являться определение вектора \mathbf{c} .

Метод чебышевской аппроксимации предполагает расчет коэффициентов c_k , входящих в состав выражения (2.33), при выполнении формулы (2.31) в частотных полосах, к которым предъявлены требования по АЧХ.

Поиск решения основан на теореме Чебышева: для того чтобы функция $R(f_0, \mathbf{c})$ являлась функцией наилучшего чебышевского приближения при заданных $K, A^*(f_0), q(f_0)$, необходимо и достаточ-

но, чтобы функция взвешенной ошибки аппроксимации

$$q(f_0) |A^*(f_0) - R(f_0, \mathbf{c})|$$

принимала наибольшие, равные друг другу по абсолютной величине и чередующиеся по знаку значения в $L + 2$ последовательно расположенных точках (точках альтернанса) интервала частот $f_0 \in [0; 0,5]$, т. е. на частотах $0 \leq f_{01} \leq f_{02} \leq \dots \leq f_{0(L+2)} \leq 0,5$.

Очевидно, что для фильтра типа 1 число точек альтернанса равно $L + 2 = (K + 3)/2$.

На рис. 2.39 показана графическая иллюстрация решения задачи чебышевской аппроксимации АЧХ КИХ-фильтра при $K = 15$. В этом случае функция $R(f_0, \mathbf{c})$ имеет 16 особых точек, в том числе 9 точек альтернанса.

2.9. Элементы проектирования цифровых фильтров с бесконечной импульсной характеристикой

Актуальность задачи синтеза рекурсивных БИХ-фильтров определяется рядом преимуществ. Так, например, приемлемая крутизна АЧХ КИХ-фильтра достигается обычно при значительном порядке фильтра ($K = 100 \dots 300$), в то время как БИХ-фильтр со сравнимыми показателями крутизны АЧХ имеет порядок 10–15. В тех случаях, когда вид ФЧХ не играет определяющей роли для практического применения разрабатываемого ЦФ, имеет смысл использовать БИХ-фильтр, что обеспечивает существенный выигрыш в быстродействии фильтра и в аппаратных затратах на его реализацию.

Решение задачи синтеза БИХ-фильтра включает в себя ряд этапов, несколько отличающийся от аналогичной последовательности операций при синтезе КИХ-фильтров.

К числу основных этапов проектирования следует отнести:

- постановку и решение задачи аппроксимации требуемых характеристик ЦФ (определение порядка фильтра и коэффициентов b_k , a_m передаточной функции);
- выбор структуры БИХ-фильтра;
- расчет разрядности входного и выходного сигналов, коэффициентов передаточной функции, кодов сигналов внутри алгоритма фильтрации;
- определение порядка расстановки звеньев в каскадных структурах с целью минимизации выходного шума;
- оценку устойчивости БИХ-фильтра;
- разработку пакета прикладных программ;
- выбор вычислителя для реализации фильтра;

- оценку характеристик созданного БИХ-фильтра.

Последний из перечисленных этапов представляет собой задачу анализа БИХ-фильтра, без выполнения которой задача синтеза не может считаться законченной.

Для решения аппроксимационной задачи используются три класса методов:

- методы преобразования аналоговых фильтров в цифровые (методы аналого-цифровой трансформации);
- прямые методы расчета;
- методы, использующие алгоритмы оптимизации.

Трудности решения задачи прямой аппроксимации желаемых характеристик дробно-рациональными передаточными функциями и достаточная развитая теория синтеза аналоговых фильтров обуславливают наиболее широкое использование методов аналого-цифровой трансформации, к которым относятся:

- метод инвариантности импульсной характеристики;
- метод билинейного преобразования;
- метод согласованного z -преобразования.

Билинейное преобразование. Синтез цифровых БИХ-фильтров на основе билинейного преобразования. Билинейное преобразование является одним из наиболее распространенных методов синтеза цифровых фильтров и обеспечивает трансформацию передаточной функции $T(p)$ аналогового фильтра-прототипа в передаточную функцию $H(z)$ цифрового фильтра.

Метод основан на замене комплексного аргумента $p = \sigma + j\omega$ комплексным аргументом $z = \alpha + j\beta$ в соответствии с правилом:

$$p = k \frac{1 - z^{-1}}{1 + z^{-1}} = k \frac{z - 1}{z + 1}, \quad (2.34)$$

что позволяет непосредственно получить передаточную функцию $H(z)$ в виде дробно-рациональной функции:

$$H(z) = T(p) \Big|_{p=k \frac{z-1}{z+1}}. \quad (2.35)$$

Из выражения (2.35) следует обратное соотношение переменных p и z :

$$z = \frac{k + p}{k - p}. \quad (2.36)$$

Коэффициент k в формулах (2.34)–(2.36) может иметь различные вещественные значения, определяемые особенностями конкретных задач синтеза цифровых фильтров. Его влияние на результат билинейного преобразования будет рассмотрено ниже.

Применение формулы (2.36) обеспечивает отображение точек комплексной p -плоскости в точки комплексной z -плоскости. В общем случае точка $p = \sigma + j\omega$ отображается в точку

$$\begin{aligned} z &= \frac{k+p}{k-p} = \frac{k+\sigma+j\omega}{k-\sigma-j\omega} = \frac{\sqrt{(k+\sigma)^2 + \omega^2} e^{j \arctg \frac{\omega}{k+\sigma}}}{\sqrt{(k-\sigma)^2 + \omega^2} e^{j \arctg \frac{-\omega}{k-\sigma}}} = \\ &= \sqrt{\frac{(k+\sigma)^2 + \omega^2}{(k-\sigma)^2 + \omega^2}} \exp \left[j \left(\arctg \frac{\omega}{k+\sigma} + \arctg \frac{\omega}{k-\sigma} \right) \right]. \end{aligned} \quad (3.37)$$

Рассмотрим на основании формулы (2.37) отображение некоторых точек p -плоскости (рис. 2.40).

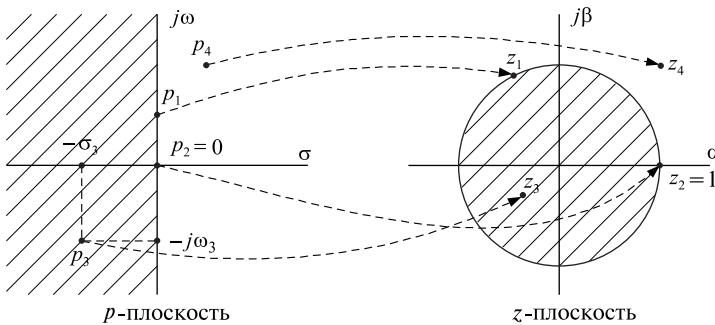


Рис. 2.40. Отображение точек при билинейном преобразовании

Точки мнимой оси p -плоскости. Пусть $p_1 = 0 + j\omega_1$ (одна из точек мнимой оси). Тогда в соответствии с формулой (2.37)

$$z_1 = 1 e^{j2 \arctg \omega_1 / k}.$$

Полученная точка z_1 имеет модуль $|z_1| = 1$ и фазовый угол $\varphi_1 = 2 \arctg(\omega_1/k)$. При монотонном изменении ω_1 от $-\infty$ до $+\infty$ фазовый угол φ_1 монотонно изменяется от $-\pi$ до $+\pi$, что означает отображение всех точек мнимой оси p -плоскости на окружность единичного радиуса в z -плоскости. Частными случаями являются отображения точки начала координат p -плоскости ($p_2 = 0$) в точку $z_2 = 1$ и точек $p = \pm j\omega$ в точку $z = -1$.

Точки левой полуплоскости p -плоскости. Пусть $p_3 = \sigma_3 - j\omega_3$, $\sigma_3 < 0$. В этом случае модуль $|z_3|$ соответствующей точки z -плоскости обязательно будет меньше 1:

$$|z_3| = \sqrt{\frac{(k+\sigma_3)^2 + \omega_3^2}{(k-\sigma_3)^2 + \omega_3^2}} < 1,$$

что означает отображение всех точек левой p -полуплоскости в точки z -плоскости, находящиеся внутри окружности единичного радиуса с центром в начале координат. Такое отображение обеспечивает соответствие условия устойчивости аналогового фильтра-прототипа, по которому все полюса p^* передаточной функции $T(p)$ находятся в левой p -полуплоскости, условию устойчивости синтезируемого цифрового фильтра, в соответствии с которым все полюса z^* цифрового фильтра должны находиться внутри единичной окружности z -плоскости.

Точки правой полуплоскости p -плоскости. Данные точки отображаются в точки z -плоскости, имеющие модуль $|z| > 1$ и лежащие за пределами единичной окружности с центром в начале координат.

Одной из основных особенностей билинейного преобразования является сопровождающая его деформация частотной шкалы. Известно, что в случае $\sigma = 0$ передаточная функция $T(p)$ аналогового фильтра соответствует его частотной характеристике $T(j\omega)$:

$$T(j\omega) = T(p) \quad \text{при } p = j\omega. \quad (2.38)$$

Частотная характеристика аналогового фильтра соответствует его передаточной функции, рассмотренной на мнимой оси p -плоскости. Соотношение, отражающее связь частотной характеристики и передаточной функции цифрового фильтра, имеет вид

$$H(e^{j\omega_{\text{ц}}T}) = H(z) \quad \text{при } z = e^{j\omega_{\text{ц}}T}, \quad (2.39)$$

т. е. частотная характеристика цифрового фильтра соответствует его передаточной функции, рассмотренной на единичной окружности z -плоскости с центром в начале координат. А поскольку мнимая ось p -плоскости при билинейном преобразовании отображается на единичную окружность z -плоскости, все максимумы и минимумы АЧХ $A(\omega)$ аналогового фильтра-прототипа сохраняются в АЧХ $A(\omega_{\text{ц}}T)$ цифрового фильтра.

Соотношения (2.38), (2.39) и зависимость (2.34) позволяют выявить связь частотных шкал аналогового фильтра-прототипа и полученного из него на основе билинейного преобразования цифрового фильтра (ω и $\omega_{\text{ц}}$ соответственно):

$$\begin{aligned} j\omega &= k \frac{e^{j\omega_{\text{ц}}T} - 1}{e^{j\omega_{\text{ц}}T} + 1} = k \frac{\cos \omega_{\text{ц}}T + j \sin \omega_{\text{ц}}T - 1}{\cos \omega_{\text{ц}}T + j \sin \omega_{\text{ц}}T + 1}; \\ j\omega - \omega \operatorname{tg} \frac{\omega_{\text{ц}}T}{2} &= jk \operatorname{tg} \frac{\omega_{\text{ц}}T}{2} - k \operatorname{tg}^2 \left(\frac{\omega_{\text{ц}}T}{2} \right); \\ \omega &= k \operatorname{tg} \frac{\omega_{\text{ц}}T}{2}. \end{aligned} \quad (2.40)$$

Деформация частотной шкалы, определяемая на основании формулы (2.40), не приводит к нарушению избирательных свойств, но смещает характеристические точки частотных характеристик по шкале частот, что необходимо учитывать при синтезе цифрового фильтра методом билинейного преобразования. Выбор различных значений коэффициента k способен управлять процессом деформации частотного диапазона. Формула (2.40) позволяет определять необходимое значение k , обеспечивающее требуемую зависимость между частотами ω и $\omega_{\text{ц}}$, и возможность синтеза АЧХ цифрового фильтра $A(\omega_{\text{ц}}T)$ по АЧХ аналогового фильтра $A(\omega)$ путем проекции точек (рис. 2.41).

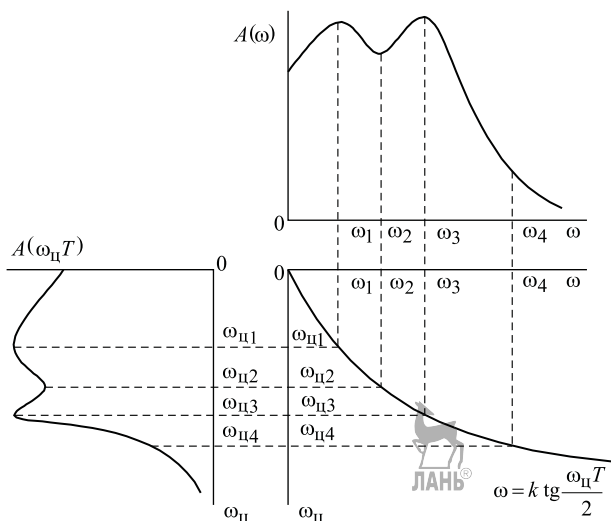


Рис. 2.41. Графическая иллюстрация использования билинейного преобразования

2.10. Специальные цифровые фильтры

Как известно, цифровой фильтр представляет собой дискретную систему, реализующую преобразование множества входных дискретных сигналов X во множество выходных сигналов Y в соответствии с оператором фильтрации $F : X \rightarrow Y$. Оператор фильтрации F определяется назначением фильтра и формируется на этапе синтеза ЦФ. Значительная часть задач цифровой фильтрации связана с обеспечением условий безыскаженной передачи сигналов в полосе пропускания фильтра (в частотной области такая задача решается равномерностью амплитудно-частотной характеристики и линейностью фазочастотной характеристики фильтра) и подавлением частотных составляющих обрабатываемого сигнала в полосе

задерживания. Традиционно такие фильтры называются *частотно-избирательными* или *селективными*. Однако существуют задачи обработки дискретных сигналов, требующие иного содержания оператора F . Соответствующие цифровые фильтры получили название *специальных* фильтров и находят широкое применение в технике связи. К числу специальных фильтров относятся фазовращатели, цифровые преобразователи Гильберта (фазовращатели на $\pi/2$), цифровые дифференциаторы и интеграторы, согласованные цифровые фильтры, медианные цифровые фильтры, ограничители мгновенных значений и т. п.

2.10.1. Цифровые преобразователи Гильберта

В теории аналоговых сигналов весьма полезным для решения ряда задач оказывается представление квазистационарного действительного сигнала $x(t)$ в виде комплексного *аналитического* сигнала:

$$\dot{x}(t) = x(t) + j\tilde{x}(t) = s(t)e^{j\varphi(t)},$$

где $s(t) = \sqrt{x^2(t) + \tilde{x}^2(t)}$ — огибающая действительного сигнала $x(t)$; $\varphi(t) = \operatorname{arctg} \frac{\tilde{x}(t)}{x(t)}$ — мгновенная фаза действительного сигнала $x(t)$.

Некоторые области использования аналитического сигнала рассмотрены в учебной дисциплине «Теория электрической связи». Особенностью аналитического сигнала $\dot{x}(t)$ среди комплексных сигналов является то, что его вещественная (действительная) $x(t)$ и мнимая $\tilde{x}(t)$ части находятся в квадратуре относительно друг друга, при этом термин «квадратура» указывает на разность фаз всех частотных составляющих сигналов $x(t)$ и $\tilde{x}(t)$ в $\pi/2$, равную четверти от полного сдвига на 2π .

Действительные сигналы $x(t)$ и $\tilde{x}(t)$ принято называть квадратурными составляющими сигнала $\dot{x}(t)$ или сигналами, сопряженными по Гильберту. Процедура получения сигнала $\tilde{x}(t)$ из известного сигнала $x(t)$ называется преобразованием Гильберта, а устройство, реализующее эту процедуру, — преобразователем Гильберта. Простейшим примером сопряженных по Гильберту сигналов являются сигналы $\cos(\omega t)$ и $\sin(\omega t)$.

В теории дискретных сигналов используется представление вещественного дискретного сигнала $x(n)$ в аналитическом виде:

$$\dot{x}(n) = x(n) + j\tilde{x}(n) = s(n)e^{j\varphi(n)}, \quad (2.41)$$

где $x(n)$ и $\tilde{x}(n)$ — действительные сигналы, сопряженные по Гильберту (отличающиеся фазами на 90°); $s(n) = \sqrt{x^2(n) + \tilde{x}^2(n)}$ — огибающая действительного сигнала $x(n)$, $\varphi(n) = \operatorname{arctg} \frac{\tilde{x}^2(n)}{x^2(n)}$ — мгновенная фаза действительного сигнала.

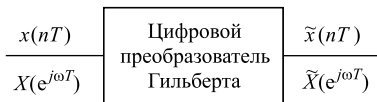


Рис. 2.42. Дискретное преобразование Гильберта

Таким образом, дискретное преобразование Гильберта (ДПГ) предназначено для формирования дискретного сигнала по известному дискретному сигналу $x(n)$. Устройство, реализующее такое преоб-

разование, называется цифровым преобразователем Гильберта (ЦПГ) или цифровым фазовращателем на 90° (в некоторых случаях к классу ЦПГ относят устройства, на выходе которых формируется пара сигналов $x^2(n)$ и $\tilde{x}^2(n)$).

Схематически процедура ДПГ представлена на рис. 2.42.

С учетом выражения (2.41) может быть определена частотная характеристика идеального ЦПГ:

$$H_r(e^{j\omega T}) = \frac{\tilde{X}(e^{j\omega T})}{X(e^{j\omega T})} = \begin{cases} -j = e^{-j\pi/2}, & 0 \leq \omega T < \pi; \\ j = e^{j\pi/2}, & \pi \leq \omega T < 2\pi. \end{cases} \quad (2.42)$$

Из выражения (2.42) легко определить амплитудно-частотную (АЧХ) и фазочастотную (ФЧХ) характеристики идеального ЦПГ ($A_r(\omega T)$ и $\varphi_r(\omega T)$ соответственно):

$$A_r(\omega T) = |H_r(e^{j\omega T})| = 1, \quad 0 \leq \omega T < 2\pi; \quad (2.43)$$

$$\varphi_r(\omega T) = \arg[H_r(e^{j\omega T})] = \begin{cases} -\pi/2, & 0 \leq \omega T < \pi; \\ \pi/2, & \pi \leq \omega T < 2\pi. \end{cases} \quad (2.44)$$

Полученные выражения (2.43) и (2.44) характеризуют частотные свойства идеального ЦПГ: амплитуды гармонических составляющих сигнала $x(n)$ при прохождении его через идеальный ЦПГ остаются неизменными, а их фазы изменяются на $\pi/2$, причем знак этого изменения зависит от частотной области (например, для основной полосы частот изменение фазы составляет $-\pi/2$).

Графическое представление АЧХ и ФЧХ идеального ЦПГ с учетом их периодичности показано на рис. 2.43 в полосе частот $0 \leq \omega T < 3\pi$.

Следует отметить, что идеальная $\varphi_r(\omega T)$, представленная на рис. 2.43, показывает лишь разницу между начальными фазами гармонических составляющих сигналов $x(n)$ и $\tilde{x}(n)$ и не учитывает время, требуемое сигналу $x(n)$ на прохождение по ЦПГ. Очевидно, что выражение (2.44) обуславливает нулевое групповое время прохождения в основной полосе частот (как, впрочем, и на других частотах), что свидетельствует о физической нереализуемости идеального ЦПГ. Кроме того, на практике невозможна реализация требуемых скачков ФЧХ на частотах $\omega T = k\pi$, $k = 0, 1, 2, 3, \dots$. Выполнимость требований ДПГ может быть реализована лишь в какой-либо рабочей полосе

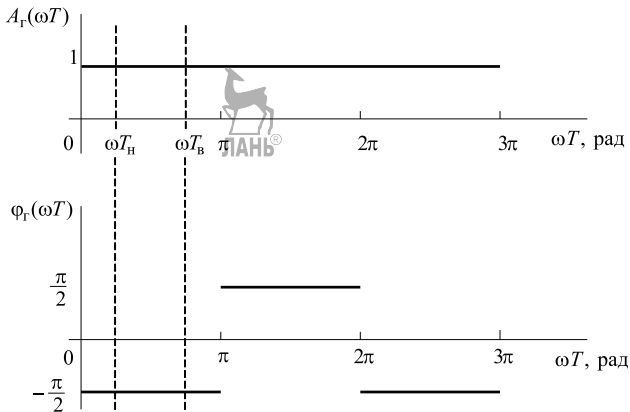


Рис. 2.43. Амплитудно-частотная и фазочастотная характеристики идеального цифрового преобразователя Гильберта

частот, которая для основной полосы частот может быть определена как $0 < \omega T_N \leq \omega T \leq \omega T_B < \pi$ (показана на рис. 2.43 пунктирными линиями). При этом АЧХ и ФЧХ физически реализуемого ЦПГ имеют следующий вид:

$$A_T(\omega T) = 1, \quad 0 < \omega T_N \leq \omega T \leq \omega T_B < \pi;$$

$$\varphi_T(\omega T) = -\frac{\pi}{2}, \quad 0 < \omega T_N \leq \omega T \leq \omega T_B < \pi.$$

Известно, что линейность ФЧХ может быть обеспечена при использовании нерекурсивных КИХ-фильтров с симметричными и антисимметричными импульсными характеристиками относительно их середины. При этом на практике чаще реализуется сдвиг фаз частотных составляющих на $\pi/2$ (что не меняет сути ДПГ) на основе КИХ-фильтров, принадлежащих к типам 3 и 4.

Один из простейших ЦПГ может быть реализован на КИХ-фильтре 2-го порядка, принадлежащем к типу 3 (фильтр четного порядка с антисимметрией импульсной характеристики). Структурная схема такого фильтра приведена на рис. 2.44.

Передаточная функция представленного фильтра определяется выражением $H(z) = b_0 - b_0 z^{-2}$. Следовательно, его частотная характеристика имеет вид

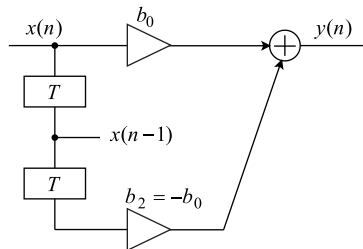


Рис. 2.44. Структурная схема ЦПГ на основе КИХ-фильтра 2-го порядка 3-го типа

$$H(e^{j\omega T}) = b_0 - b_0 e^{-2j\omega T} = b_0 - b_0 \cos 2\omega T + j b_0 \sin 2\omega T.$$

Фазочастотная характеристика исследуемого фильтра имеет вид

$$\begin{aligned}\varphi(\omega T) &= \operatorname{arctg} \frac{b_0 \sin 2\omega T}{b_0 - b_0 \cos 2\omega T} = \operatorname{arctg} \frac{\sin 2\omega T}{1 - \cos 2\omega T} = \\ &= \operatorname{arctg} \frac{2 \sin \omega T \cos \omega T}{2 \sin^2 \omega T} = \operatorname{arctg}(\operatorname{ctg} \omega T) = \\ &= \operatorname{arctg}[\operatorname{tg}(\pi/2 - \omega T)] = \pi/2 - \omega T.\end{aligned}\quad (2.45)$$

Графическое представление рассчитанной $\varphi(\omega T)$ для основной полосы частот показано на рис. 2.45 (верхняя наклонная прямая).

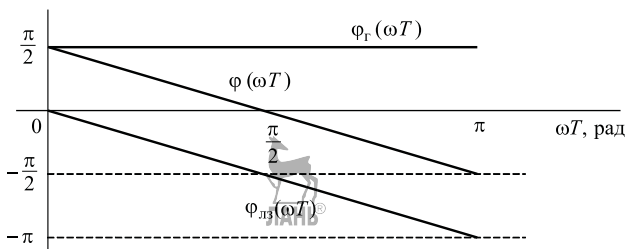


Рис. 2.45. Фазочастотные характеристики различных участков КИХ-фильтра 2-го порядка 3-го типа

При анализе выражения (2.45) следует принимать во внимание, что исследуемая ФЧХ содержит два компонента:

- слагаемое $-\omega T$, определяющее постоянное групповое время прохождения $\tau(\omega T) = -\frac{d(-\omega T)}{d\omega}$, обуславливающее одинаковый набег фаз всех гармонических составляющих сигнала $x(n)$, но не вносящее изменений в их соотношение;
- слагаемое $\pi/2$, обеспечивающее требуемое приращение фаз всех гармонических составляющих сигнала $x(n)$.

Таким образом, анализ ФЧХ (2.45) свидетельствует о том, что сигнал $y(n)$, формируемый на выходе исследуемого фильтра (см. рис. 2.45), сопряжен по Гильберту не с входным сигналом $x(n)$, а с его копией $x(n-1)$, задержанной на период T . Именно для этих двух сигналов обеспечивается сдвиг начальных фаз всех гармонических составляющих на $\pi/2$. Доказательство этого представлено ниже.

Фазовый спектр $\varphi_{y(n)}(\omega T)$ выходного сигнала $y(n)$ определяется суммой фазового спектра $\varphi_{x(n)}(\omega T)$ входного сигнала $x(n)$ и рассчитанной ФЧХ $\varphi(\omega T)$ цифрового фильтра:

$$\varphi_{y(n)}(\omega T) = \varphi_{x(n)}(\omega T) + \varphi(\omega T).$$

Фазовый спектр $\varphi_{x(n-1)}(\omega T)$ сигнала $x(n-1)$ определяется суммой фазового спектра $\varphi_{x(n)}(\omega T)$ входного сигнала $x(n)$ и ФЧХ

$\varphi_{\text{ЛЗ}}(\omega T)$ линии задержки на период дискретизации T :

$$\varphi_{x(n-1)}(\omega T) = \varphi_{x(n)}(\omega T) + \varphi_{\text{ЛЗ}}(\omega T).$$

Известно, что частотная характеристика линии задержки на период T определяется выражением $H_{\text{ЛЗ}}(e^{j\omega T}) = e^{-j\omega T}$, из чего следует, что ее ФЧХ $\varphi_{\text{ЛЗ}}(\omega T) = -\omega T$ (нижняя наклонная прямая на рис. 2.45).

Фазочастотная характеристика участка исследуемого фильтра между сигналами $x(n-1)$ и $y(n)$ определяется разностью их фазовых спектров:

$$\begin{aligned} \varphi_{\Gamma}(\omega T) &= \varphi_{y(n)}(\omega T) - \varphi_{x(n-1)}(\omega T) = \\ &= \varphi(\omega T) - \varphi_{\text{ЛЗ}}(\omega T) = \frac{\pi}{2} - \omega T + \omega T = \frac{\pi}{2} \end{aligned}$$

и полностью удовлетворяет требованиям, предъявляемым к ФЧХ цифрового преобразователя Гильберта. График $\varphi_{\Gamma}(\omega T)$ показан на рис. 2.45 (верхняя прямая линия).

Таким образом, для рассмотренного ЦПГ элемент задержки на период дискретизации T , на выходе которого формируется сигнал $x(n-1)$, выполняет роль согласующей линии задержки (СЛЗ). Существенно, что аналогичные СЛЗ можно выделить в любых ЦПГ, построенных на основе КИХ-фильтров 3-го и 4-го типов.

АЧХ исследуемого фильтра определяется выражением

$$\begin{aligned} A_{\Gamma}(\omega T) &= \sqrt{(b_0 - b_0 \cos 2\omega T)^2 + (b_0 \sin 2\omega T)^2} = \\ &= \sqrt{b_0^2 - 2b_0^2 \cos 2\omega T + b_0^2 \cos^2 2\omega T + b_0^2 \sin^2 2\omega T} = \\ &= \sqrt{b_0^2(2 - 2 \cos 2\omega T)} = \sqrt{b_0^2 4 \sin^2 \omega T} = 2|b_0 \sin \omega T|. \end{aligned}$$

Очевидно, расчет АЧХ участка исследуемого фильтра между сигналами $x(n-1)$ и $y(n)$ даст такой же результат, ибо СЛЗ не вносит амплитудных изменений в обрабатываемые дискретные сигналы.

Графическое представление рассчитанной АЧХ показано на рис. 2.46.

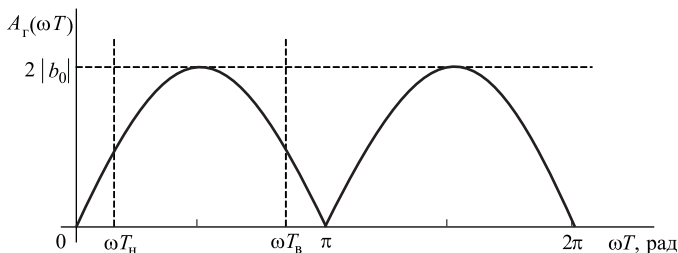


Рис. 2.46. АЧХ ЦПГ, реализованного на КИХ-фильтре 2-го порядка 3-го типа

Очевидно, что реальная АЧХ далека от АЧХ идеального ЦПГ. Это обусловлено низким порядком использованного КИХ-фильтра и тем фактом, что АЧХ всех КИХ-фильтров 3-го типа на частотах, кратных π , равны нулю. Практика показывает, что достижение АЧХ, близкой к идеальной в рабочей полосе частот $0 < \omega T_n \leq \omega T \leq \omega T_v < \pi$, обеспечивается при использовании КИХ-фильтров 3-го и 4-го типов примерно 20-го порядка, синтезируемых на основе апробированных методик.

Примером использования ЦПГ в системах связи является процедура демодуляции однополосного сигнала, получаемого путем выделения одной из боковых полос амплитудно-модулированного сигнала. В этом случае формирование сопряженных по Гильберту сигналов $x(n)$ и $\tilde{x}(n)$ (в рассмотренном примере $x(n-1)$ и $y(n)$) предшествует операции вычисления огибающей принятого сигнала $x(n)$.

2.10.2. Цифровые дифференциаторы и интеграторы

Из теории аналоговых сигналов известно, что дифференциатором называется устройство, выполняющее операцию дифференцирования, в результате чего по известному сигналу $x(t)$ формируется сигнал $y(t) = k \frac{dx(t)}{dt}$, где k — масштабирующий коэффициент, выбранный в соответствии с условиями решаемой задачи. Естественно, что указанная операция осуществляется в требуемом частотном диапазоне, называемом рабочей областью дифференциатора. Наиболее показательным примером аналогового дифференцирования является преобразование синусоидального сигнала:

$$y(t) = k \frac{d \sin(\omega t)}{dt} = k\omega \cos(\omega t).$$

Переход в область дискретных сигналов приводит рассмотренный пример к виду, показанному на рис. 2.47.

$x(nT) = \sin(\omega nT)$	Цифровой дифференциатор $H_d(z)$	$y(nT) = k \frac{\omega T}{T} \cos(\omega nT)$
$X(z) = Z[\sin(\omega nT)]$		$Y(z) = Z\left[k \frac{\omega T}{T} \cos(\omega nT)\right]$

Рис. 2.47. Схематическое представление процедуры цифрового дифференцирования дискретной синусоиды

Для определения частотной характеристики цифрового дифференциатора (ЦД) воспользуемся расчетом его передаточной функции $H_d(z)$ с учетом известных z -образов гармонических сигналов:

$$Z[\sin(\omega nT)] = \frac{\sin(\omega nT)z^{-1}}{1 - 2\cos(\omega nT)z^{-1} + z^{-2}};$$

$$Z[\cos(\omega nT)] = \frac{1 - \cos(\omega nT)z^{-1}}{1 - 2\cos(\omega nT)z^{-1} + z^{-2}}.$$

Следовательно,

$$H_d(z) = \frac{Y(z)}{X(z)} = k \frac{\omega T}{T} \frac{1 - \cos(\omega nT)z^{-1}}{\sin(\omega nT)z^{-1}}. \quad (2.46)$$

Выражение (2.46) с учетом известных соотношений $z = e^{j\omega T}$, $\cos(\omega T) = \frac{e^{j\omega T} + e^{-j\omega T}}{2}$ и $\sin(\omega T) = \frac{e^{j\omega T} - e^{-j\omega T}}{2j}$ позволяет получить искомую ЧХ идеального цифрового дифференциатора:

$$H_d(e^{j\omega T}) = jk\omega = k\omega e^{j\pi/2}. \quad (2.47)$$

Анализ выражения (2.47) позволяет сделать вывод о том, что ЧХ идеального ЦД является линейной и чисто мнимой функцией частоты. При определении коэффициента $k = T/\pi$ частотная характеристика приобретает вид $H_d(e^{j\omega T}) = j\omega T/\pi$, откуда следует выражение для АЧХ: $A_d(\omega T) = \omega T/\pi$. Очевидно, что в основной полосе частот АЧХ линейно изменяется от 0 до 1 (рис. 2.48).

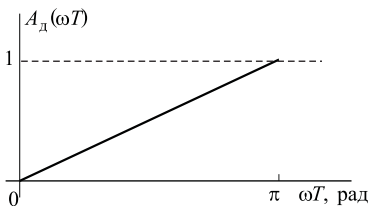


Рис. 2.48. АЧХ идеального цифрового дифференциатора

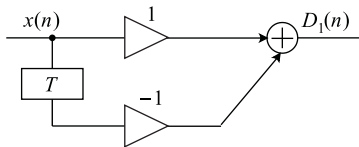


Рис. 2.49. ЦД на основе КИХ-фильтра 1-го порядка 4-го типа

Требуемая характеристика на практике может быть обеспечена на основе КИХ-фильтров с линейной ФЧХ 4-го типа (в основной полосе частот $0 \leq \omega T \leq \pi$) или 3-го типа (в рабочей полосе частот $0 \leq \omega T \leq \omega T_{\text{раб}} < \pi$, так как АЧХ фильтров данного типа на частоте $\omega T = \pi$ равна нулю).

В качестве примера рассмотрим простейший ЦД на основе КИХ-фильтра 1-го порядка 4-го типа (рис. 2.49).

Представленный ЦД вычисляет разность 1-го порядка $D_1(nT) = x(nT) - x(nT - T)$, являющуюся дискретным аналогом производной 1-го порядка.

Очевидно, что передаточная функция исследуемого ЦД имеет вид $H(z) = 1 - z^{-1}$, а его частотная характеристика определяется выражением

$$H(e^{j\omega T}) = 1 - e^{-j\omega T} = (e^{j\omega T/2} - e^{-j\omega T/2})e^{-j\omega T/2} = 2\sin(\omega T/2)e^{-j\omega T/2},$$

модуль которого представляет собой АЧХ ЦД:

$$A(\omega T) |H(e^{j\omega T})| = 2 |\sin(\omega T/2)|.$$

Графическое представление рассчитанной АЧХ показано на рис. 2.50. Для лучшего приближения реальной АЧХ к линейной в основной полосе частот необходимо использовать КИХ-фильтры 4-го типа, имеющие больший порядок.

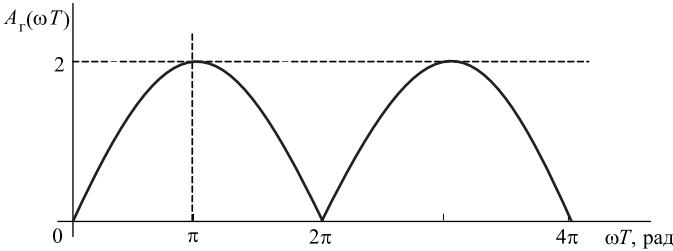


Рис. 2.50. АЧХ ЦД на основе КИХ-фильтра 1-го порядка 4-го типа

Цифровой интегратор в общем случае выполняет операцию, обратную дифференцированию (с точностью до постоянного слагаемого). Поэтому простейший цифровой интегратор (ЦИ) должен иметь передаточную функцию, обратную передаточной функции ЦД:

$$H(z) = (1 - z^{-1})^{-1} = \frac{1}{1 - z^{-1}}.$$

Такая ПФ обеспечивается рекурсивным ЦФ 1-го порядка, структурная схема которого показана на рис. 2.51.

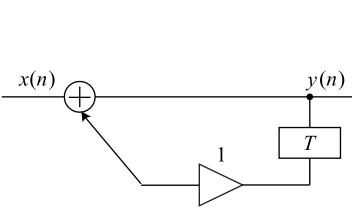


Рис. 2.51. ЦИ на основе рекурсивного ЦФ 1-го порядка

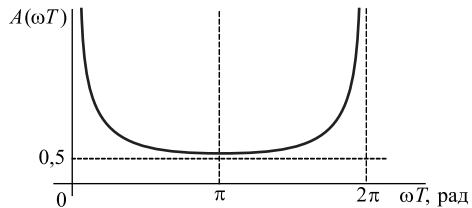


Рис. 2.52. АЧХ ЦИ на основе рекурсивного ЦФ 1-го порядка

Частотная характеристика определяется выражением

$$H(e^{j\omega T}) = [2 \sin(\frac{\omega T}{2})]^{-1} e^{j\frac{\omega T}{2}},$$

модуль которого представляет собой АЧХ ЦИ:

$$A(\omega T) = |H(e^{j\omega T})| = [2 |\sin(\frac{\omega T}{2})|]^{-1}.$$

Графическое представление рассчитанной АЧХ показано на рис. 2.52. Для лучшего приближения реальной АЧХ к линейной в основной полосе частот необходимо использовать БИХ-фильтры большего порядка.

Цифровые дифференциаторы и интеграторы широко применяются в системах автоматического управления, оптимальных приемниках и других элементах телекоммуникационных систем.

2.10.3. Цифровые согласованные КИХ-фильтры

В системах инфокоммуникаций известна задача принятия решения о форме принятого сигнала при известном конечном множестве возможных. Такая задача решается с помощью согласованных фильтров, импульсная характеристика которых формируется с учетом формы конкретного распознаваемого сигнала.

В системах цифровой связи используются цифровые согласованные фильтры (СФ), определяемые следующим образом: фильтром, согласованным с финитным дискретным сигналом $x(nT)$ длительностью n_0T , называется фильтр с импульсной характеристикой $h_c(nT)$, удовлетворяющей условию $h_c(nT) = \alpha T x(n_0T - nT)$, где α — масштабирующий коэффициент, выбираемый с учетом энергии распознаваемого сигнала $x(nT)$.

В случае распознавания сигналов с равными энергиями $\alpha T = 1$ требование к импульсной характеристике согласованного фильтра принимает следующий вид:

$$h_c(nT) = x(n_0T - nT). \quad (2.48)$$

Из выражения (2.48) следует, что импульсная характеристика СФ $h_c(nT)$ должна повторять отсчеты распознаваемого сигнала $x(nT)$ в обратном порядке (часто используется термин «в зеркальном отображении») (рис. 2.53). Очевидно, что реализация СФ возможна на нерекурсивных ЦФ.

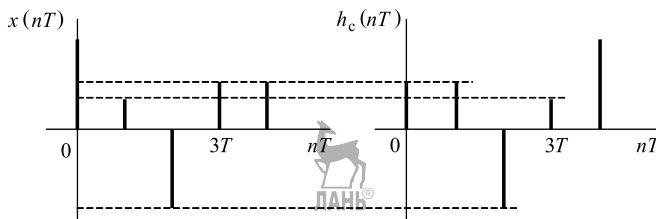


Рис. 2.53. Финитный дискретный сигнал и импульсная характеристика согласованного с ним фильтра для случая $n_0 = 4$

Частотная характеристика СФ имеет вид

$$H_c(e^{j\omega T}) = X(e^{-j\omega T})e^{-j\omega n_0 T}, \quad (2.49)$$

где $X(e^{-j\omega T})$ — комплексно-сопряженный спектр относительно спектра $X(e^{j\omega T})$ сигнала $x(nT)$.

Анализ ЧХ (2.49) позволяет сделать следующие выводы:

- амплитудно-частотная характеристика фильтра, согласованного с сигналом $x(nT)$, совпадает (при $\alpha T = 1$) с амплитудным спектром сигнала $x(nT)$:

$$A_c(\omega T) = |H_c(e^{j\omega T})| = |X(e^{j\omega T})|;$$

- фазочастотная характеристика СФ

$$\varphi_c(\omega T) = -\{\arg[X(e^{j\omega T})] + \omega n_0 T\}$$

компенсирует фазовые сдвиги всех спектральных составляющих ожидаемого сигнала, вследствие чего спектральные составляющие реакции СФ оказываются синфазными и, суммируясь, дают максимум реакции на выходе фильтра в момент окончания входного сигнала $nT = n_0 T$. Обязательным условием корректного функционирования обнаружителя сигналов на СФ является полная синхронизация передатчика и приемника сигналов.

На рис. 2.54 показана структурная схема обнаружителя сигналов на цифровых согласованных фильтрах для бинарного случая.

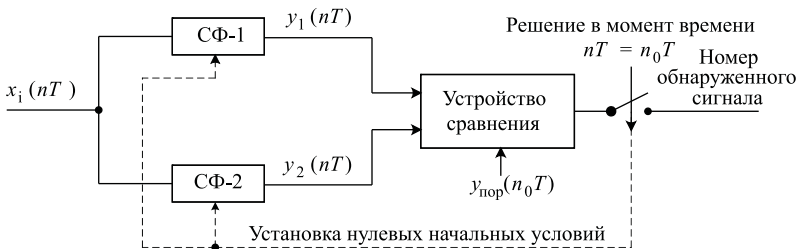


Рис. 2.54. Структурная схема обнаружителя сигналов на двух СФ

Входной сигнал $x_i(nT)$, $i = 1, 2$, поступает одновременно на оба СФ, каждый из которых согласован с одним из двух возможных сигналов. В теории ЦОС доказано, что в момент окончания сигнала $x_i(nT)$ (в момент времени $n_0 T$) максимальное значение реакции будет наблюдаться на выходе того СФ, который согласован с данным сигналом:

$$y_i(n_0 T) = \sum_{k=0}^{n_0} h_{ci}(kT) x_i(n_0 T - kT).$$

Сравнение значений реакций $y_i(n_0T)$ выполняется устройством сравнения, которое и формирует решение о принятом сигнале $x_i(nT)$. Достаточно часто для принятия решения в канале с помехами используется сравнение максимального значения реакции $y_i(n_0T)$ с пороговым значением $y_{\text{пор}}(n_0T)$. После принятия решения в момент времени $nT = n_0T$ в согласованных фильтрах должны быть обеспечены нулевые начальные условия.

Ниже рассмотрено применение обнаружителя сигналов, показанного на рис. 2.54, для общего бинарного случая при идеальном канале связи. Ансамбль входных дискретных сигналов сформирован из противоположных аналоговых сигналов (например сигналов двоичной фазовой модуляции со сдвигом начальной фазы на π): $x_1(nT) = (0,1; -0,4; 0,3)$, $x_2(nT) = (-0,1; 0,4; -0,3)$. Импульсные характеристики $h_{ci}(nT)$ КИХ-фильтров СФ-1 и СФ-2, согласованных с сигналами $x_i(nT)$, имеют вид $h_{c1}(nT) = (0,3; -0,4; 0,1)$, $h_{c2}(nT) = (-0,3; 0,4; -0,1)$.

Пусть на вход обнаружителя с выхода идеального канала связи поступает сигнал $x_1(nT)$. При этом на выходе фильтров формируются реакции, определяемые периодической сверткой:

$$y_i(nT) = \sum_{k=0}^2 h_{ci}(kT)x_1(nT - kT).$$

В момент времени $n_0 = 2T$ реакции на выходе фильтров имеют следующие значения:

$$\begin{aligned} y_1(2T) &= h_{c1}(0)x_1(2T) + h_{c1}(T)x_1(T) + h_{c1}(2T)x_1(0) = \\ &= 0,3 \cdot 0,3 + (-0,4)(-0,4) + 0,1 \cdot 0,1 = 0,26; \\ y_2(2T) &= h_{c2}(0)x_1(2T) + h_{c2}(T)x_1(T) + h_{c2}(2T)x_1(0) = \\ &= (-0,3) \cdot 0,3 + 0,4(-0,4) + (-0,1) \cdot 0,1 = -0,26. \end{aligned}$$

Следовательно, устройство сравнения на основе неравенства $y_1(2T) > y_2(2T)$ принимает решение о приеме сигнала $x_1(nT)$.

Следует иметь в виду, что в общем случае достоверность обнаружения сигналов на основе цифровых согласованных фильтров зависит от многих факторов, в первую очередь — от степени непохожести сигналов и помеховой обстановки в канале связи.

2.11. Собственные шумы цифровых устройств

Частные шумы квантования, возникающие на выходах операционных узлов цифровой системы (процессора) в результате квантования результатов арифметических операций, формируют собственный шум системы, рассматриваемый, как правило, в точке выхода

системы ЦОС. Причина его возникновения заключается в переходе от большей разрядности операционных узлов вычислителя к меньшей разрядности ячеек памяти. В общем случае анализ собственного шума гораздо сложнее, чем анализ эффектов квантования входного и выходного сигналов. Для такого анализа необходимо знать:

- точки системы ЦОС, в которых выполняется квантование результатов арифметических операций;
- статистические характеристики (в наиболее простом случае — математическое ожидание и дисперсию) частных шумовых сигналов, генерируемых отдельными узлами квантования;
- путь, который проходит каждый шумовой сигнал от своего источника (узла квантования) до выхода из системы (процессора); очевидно, что эти пути определяются структурой системы ЦОС;
- архитектуру и параметры цифрового устройства (процессора), на котором реализована система.

В частном случае реализации процедуры цифровой фильтрации собственный шум системы ЦОС естественно определить как собственный шум цифрового фильтра, причиной появления которого является необходимость хранения результатов суммирования, представленных в формате m_{acc} -разрядных выходных регистров сумматоров, в $m_{\text{п}}$ -разрядных ячейках памяти с учетом неравенства $m_{\text{acc}} > m_{\text{п}}$. Линейная модель формирования собственного шума $\epsilon_{\text{вых}}^c(nT)$ цифрового фильтра, содержащего L узлов квантования результатов суммирования, представлена на рис. 2.55.

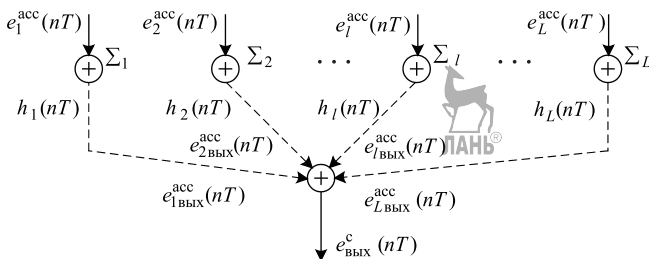


Рис. 2.55. Линейная модель формирования собственного шума цифрового фильтра

Каждый из частных шумов квантования $e_i^{\text{acc}}(nT)$, характеризующийся математическим ожиданием μ_i^{acc} и дисперсией $(\sigma_i^{\text{acc}})^2$, проходит свой путь до точки выхода из системы. Важными для дальнейших расчетов характеристиками частных путей являются их импульсные характеристики $h_i(nT)$ и значения АЧХ этих путей на нулевой частоте $A_i(0)$.

Из теории прохождения случайных процессов по линейным системам известны выражения для расчета статистических характеристик частных шумовых сигналов, приведенных к точке выхода из системы $-e_{l\text{вых}}^{\text{acc}}(nT)$:

- математическое ожидание $\mu_{l\text{вых}}^{\text{acc}} = \mu_l^{\text{acc}} A_l(0)$;
- дисперсия $(\sigma_{l\text{вых}}^{\text{acc}})^2 = (\sigma_l^{\text{acc}})^2 \sum_{n=0}^{\infty} h^2(nT)$.

В условиях статистической независимости частных шумовых сигналов $e_{l\text{вых}}^{\text{acc}}(nT)$ собственный шум цифрового фильтра в точке выхода определяется суммой частных шумовых сигналов:

$$e_{\text{вых}}^c(nT) = \sum_{l=1}^L e_{l\text{вых}}^{\text{acc}}(nT)$$

и характеризуется математическим ожиданием $\mu_{\text{вых}}^c = \sum_{l=1}^L \mu_{l\text{вых}}^{\text{acc}}$ и дисперсией $(\sigma_{\text{вых}}^c)^2 = \sum_{l=1}^L (\sigma_{l\text{вых}}^{\text{acc}})^2$.

Пример. Исследуемая система цифровой обработки сигналов, функциональная схема которой представлена на рис. 2.56, включает в себя следующие устройства:

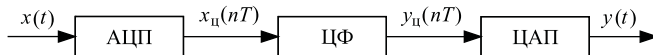


Рис. 2.56. Функциональная схема исследуемой цифровой системы

- 16-разрядный АЦП ($m_{\text{АЦП}} = 16$) и 8-разрядный ЦАП ($m_{\text{ЦАП}} = 8$), использующие представление чисел с фиксированной точкой в дополнительном коде; используемый способ квантования — округление;
- цифровой фильтр (ЦФ), реализованный на цифровом процессоре обработки сигналов (ЦПОС) с 16-разрядной памятью ($m_n = 16$), имеющий передаточную функцию

$$H(z) = \frac{0,1 + 0,6z^{-1}}{1 + 0,3z^{-1}}$$

и построенный по канонической структуре 1-го типа.

Требуется рассчитать полный выходной шум квантования сигналов в исследуемой системе.

Передаточная функция фильтра позволяет классифицировать его как рекурсивный и определить линейно-разностное уравнение, характеризующее закон фильтрации сигнала во временной области:

$$y_{\text{ц}}(nT) = 0,1x_{\text{ц}}(nT) + 0,6x_{\text{ц}}(nT - T) - 0,3y_{\text{ц}}(nT - T).$$

Каноническая структура 1-го типа для рекурсивного цифрового фильтра, соответствующего полученному ЛРУ, представлена на рис. 2.57.

Импульсная характеристика анализируемого РЦФ в общем виде определяется выражением

$$h(nT) = \begin{cases} b_0, & n = 0; \\ b_0(-a_1)^n + b_1(-a_1)^{n-1}, & n > 0, \end{cases}$$

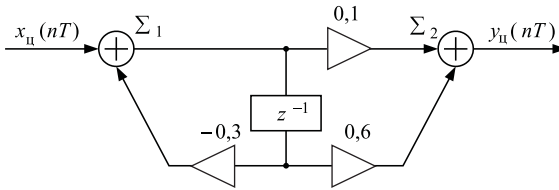


Рис. 2.57. РЦФ с канонической структурой 1-го типа

что с учетом коэффициентов фильтра приводит к виду

$$h(nT) = \begin{cases} 0,1, & n = 0; \\ 0,1(-0,3)^n + 0,6(-0,3)^{n-1}, & n > 0. \end{cases}$$

Процесс цифровой обработки сигналов в исследуемой системе характеризуется наличием в точке выхода процессора (цифрового фильтра) суммарного шума квантования $e_{\text{вых}}(nT)$, обусловленного шумом квантования входного сигнала, собственным шумом процессора (цифрового фильтра) и шумом квантования выходного сигнала.

На рис. 2.58 представлена линейная шумовая модель системы, отображающая точки возникновения и обозначения частных шумовых сигналов.

Представленная линейная шумовая модель исследуемой цифровой системы и теоретические положения, изложенные выше, позволяют определить характеристики шума квантования входного сигнала $e^{\text{АЦП}}(nT)$.

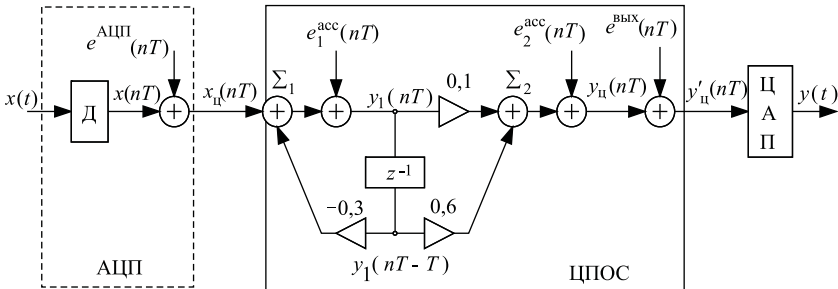


Рис. 2.58. Линейная шумовая модель исследуемой цифровой системы: $x(t)$ — исходный аналоговый сигнал; $x(nT)$ — дискретный сигнал на выходе дискретизатора; $x_u(nT)$ — цифровой $m_{\text{АЦП}}$ -разрядный сигнал на выходе АЦП; $y_1(nT)$ — промежуточный цифровой m_n -разрядный сигнал цифрового фильтра; $y_u(nT)$ — выходной цифровой m_n -разрядный сигнал цифрового фильтра; $y'_u(nT)$ — выходной цифровой $m_{\text{ЦАП}}$ -разрядный сигнал цифрового фильтра; $y(t)$ — аналоговый сигнал на выходе системы; $e^{\text{АЦП}}(nT)$ — дискретный шумовой сигнал, соответствующий шуму квантования в АЦП и определяемый заменой дискретного сигнала $x(nT)$ цифровым сигналом $x_u(nT)$ в виде $m_{\text{АЦП}}$ -разрядного выходного кода АЦП; $e_1^{\text{acc}}(nT), e_2^{\text{acc}}(nT)$ — дискретные шумовые сигналы, соответствующие шумам квантования сигналов на выходе сумматоров Σ_1 и Σ_2 цифрового фильтра при переписывании этих сигналов из m_{acc} -разрядных выходных регистров сумматоров в m_n -разрядное устройство памяти (собственный шум цифрового фильтра); $e^{\text{вых}}(nT)$ — дискретный шумовой сигнал, соответствующий шуму квантования выходного m_n -разрядного сигнала ЦФ $y_u(nT)$ при формировании цифрового $m_{\text{ЦАП}}$ -разрядного сигнала $y'_u(nT)$, поступающего на вход ЦАП

При равномерном квантовании способом округления сигнала $x(nT)$ в сигнал $x_{\Pi}(nT)$, значения которого представлены в форме с фиксированной точкой в дополнительном коде, характеристики сигнала $e^{\text{АПП}}(nT)$ определяются следующим образом:

- математическое ожидание сигнала $e^{\text{АПП}}(nT)$

$$\mu^{\text{АПП}} = 0;$$

- дисперсия сигнала $e^{\text{АПП}}(nT)$

$$(\sigma^{\text{АПП}})^2 = \frac{Q^2}{12},$$

где $Q = 2^{-(m_{\text{АПП}}-1)}$ — шаг квантования, определяемый весом младшего числового разряда. Тогда

$$(\sigma^{\text{АПП}})^2 = \frac{1}{12 \cdot 2^{30}} = 7,761 \cdot 10^{-11}$$

или, в логарифмических единицах,

$$(\sigma^{\text{АПП}})^2, \text{ дБ} = 10 \lg(\sigma^{\text{АПП}})^2 = -101,1 \text{ дБ}.$$

Кроме шума квантования входного сигнала, рассмотренного выше, на функционирование исследуемой системы оказывают влияние следующие шумовые сигналы:

- собственный шум процессора (цифрового фильтра), компонентами которого являются частные шумовые сигналы $e_1^{\text{acc}}(nT)$ и $e_2^{\text{acc}}(nT)$, определяемые процедурами переквантования сигналов внутри ЦПОС при условии $m_n < m_{\text{acc}}$;
- шум квантования выходного сигнала $e^{\text{ВЫХ}}(nT)$, определяемый процедурой переквантования сигналов на стыке ЦПОС и ЦАП при условии $m_n > m_{\text{ЦАП}}$; этот вид шума при анализе, как правило, относится к точке выхода ЦПОС.

1. Условия формирования шумовых сигналов $e_1^{\text{acc}}(nT)$ и $e_2^{\text{acc}}(nT)$ идентичны. При переписывании сигналов из выходных регистров сумматоров в m_n -разрядное устройство памяти используется форма представления чисел с фиксированной точкой и квантование способом округления. Следовательно, математические ожидания обоих сигналов равны нулю, а одинаковые дисперсии составляют

$$(\sigma_1^{\text{acc}})^2 = (\sigma_2^{\text{acc}})^2 = \frac{Q^2}{12} = \frac{2^{-30}}{12} = 7,761 \cdot 10^{-11}$$

или, в логарифмических единицах,

$$(\sigma_1^{\text{acc}})^2, \text{ дБ} = (\sigma_2^{\text{acc}})^2, \text{ дБ} = 10 \lg(7,761 \cdot 10^{-11}) = -101,1 \text{ дБ}.$$

2. При равномерном квантовании способом округления и представлении чисел в форме с фиксированной точкой в дополнительном коде характеристики $e^{\text{ВЫХ}}(nT)$ полностью соответствуют характеристикам $e^{\text{ЦАП}}(nT)$ и составляют:

- математическое ожидание сигнала $e^{\text{ВЫХ}}(nT)$

$$\mu^{\text{ВЫХ}} = \mu^{\text{ЦАП}} = 0;$$

- дисперсия сигнала $e^{\text{ВЫХ}}(nT)$

$$(\sigma^{\text{ВЫХ}})^2 = (\sigma^{\text{ЦАП}})^2 = \frac{Q^2}{12} = \frac{1}{12 \cdot 2^{14}} = 5,086 \cdot 10^{-6}$$

или, в логарифмических единицах,

$$(\sigma^{\text{ВЫХ}})^2, \text{ дБ} = 10 \lg(\sigma^{\text{ЦАП}})^2 = -52,94 \text{ дБ}.$$

Очевидно, что полный выходной шум квантования $e^{\text{ВЫХ}}(nT)$, возникающий в исследуемой системе и отнесенный к точке выхода ЦПОС, определяется влиянием в этой точке рассмотренных выше шумовых сигналов $e^{\text{АПП}}(nT)$, $e_1^{\text{acc}}(nT)$, $e_2^{\text{acc}}(nT)$ и $e^{\text{ВЫХ}}(nT)$. Правомерное допущение о независимости указанных частных шумовых

сигналов позволяет определить $e_{\text{вых}}(nT)$ суммой реакций системы (в точке выхода ЦПОС) на шумы $e^{\text{АПП}}(nT)$, $e_1^{\text{acc}}(nT)$, $e_2^{\text{acc}}(nT)$ и $e^{\text{вых}}(nT)$:

$$e_{\text{вых}}(nT) = e_{\text{вых}}^{\text{АПП}}(nT) + e_{1\text{вых}}^{\text{acc}}(nT) + e_{2\text{вых}}^{\text{acc}}(nT) + e_{\text{вых}}^{\text{вых}}(nT). \quad (2.50)$$

Существенно, что в полученном выражении сумма $e_{1\text{вых}}^{\text{acc}}(nT) + e_{2\text{вых}}^{\text{acc}}(nT)$ определяют собственный шум процессора.

Математическое ожидание и дисперсия выходного шума $e_{\text{вых}}(nT)$ определяются в виде сумм математических ожиданий и дисперсий соответствующих составляющих:

$$\begin{aligned} \mu_{\text{вых}} &= \mu_{\text{вых}}^{\text{АПП}} + \mu_{1\text{вых}}^{\text{acc}} + \mu_{2\text{вых}}^{\text{acc}} + \mu_{\text{вых}}^{\text{вых}}; \\ (\sigma_{\text{вых}})^2 &= (\sigma_{\text{вых}}^{\text{АПП}})^2 + (\sigma_{1\text{вых}}^{\text{acc}})^2 + (\sigma_{2\text{вых}}^{\text{acc}})^2 + (\sigma_{\text{вых}}^{\text{вых}})^2. \end{aligned}$$

Ниже рассмотрены слагаемые выражения (2.50).

1. Анализ составляющей $e_{\text{вых}}^{\text{АПП}}(nT)$.

Шумовой сигнал $e_{\text{вых}}^{\text{АПП}}(nT)$ является реакцией исследуемой системы ЦОС (в точке выхода ЦПОС) на шумовой сигнал $e^{\text{АПП}}(nT)$. Так как сигнал $e^{\text{АПП}}(nT)$ проходит в точку выхода ЦПОС через весь фильтр, то на основании закономерностей прохождения случайных сигналов через линейную систему можно записать выражения для оценки статистических характеристик $e_{\text{вых}}^{\text{АПП}}(nT)$:

- математическое ожидание сигнала $e_{\text{вых}}^{\text{АПП}}(nT)$

$$\mu_{\text{вых}}^{\text{АПП}} = 0;$$
- дисперсия сигнала $e_{\text{вых}}^{\text{АПП}}(nT)$:

$$(\sigma_{\text{вых}}^{\text{АПП}})^2 = (\sigma^{\text{АПП}})^2 \sum_{n=0}^{\infty} h^2(nT),$$

где

$$\begin{aligned} \sum_{n=0}^{\infty} h^2(nT) &= b_0^2 + \sum_{n=1}^{\infty} [b_0(-a_1)^n + b_1(-a_1)^{n-1}]^2 = \\ &= \frac{b_0^2 - 2b_0b_1a_1 + b_1^2}{1 - a_1^2} = \frac{0,334}{0,91} \approx 0,367. \end{aligned}$$

Тогда

$$(\sigma_{\text{вых}}^{\text{АПП}})^2 = 7,761 \cdot 10^{-11} \cdot 0,367 = 2,848 \cdot 10^{-11}$$

или, в логарифмических единицах,

$$(\sigma_{\text{вых}}^{\text{АПП}})^2, \text{ дБ} = 10 \lg(\sigma_{\text{вых}}^{\text{АПП}})^2 = -105,45 \text{ дБ}.$$

2. Анализ составляющей $e_{1\text{вых}}^{\text{acc}}(nT)$.

Шумовой сигнал $e_{1\text{вых}}^{\text{acc}}(nT)$ является реакцией исследуемой системы ЦОС (в точке выхода ЦПОС) на шумовой сигнал $e_1^{\text{acc}}(nT)$. Так как сигнал $e_1^{\text{acc}}(nT)$ проходит к точке выхода ЦПОС через весь фильтр, то характеристики $e_{1\text{вых}}^{\text{acc}}(nT)$ оцениваются следующим образом:

- математическое ожидание сигнала $e_{1\text{вых}}^{\text{acc}}(nT)$

$$\mu_{1\text{вых}}^{\text{acc}} = 0;$$
- дисперсия сигнала $e_{1\text{вых}}^{\text{acc}}(nT)$

$$(\sigma_{1\text{вых}}^{\text{acc}})^2 = (\sigma_1^{\text{acc}})^2 \sum_{n=0}^{\infty} h^2(n) = 7,761 \cdot 10^{-11} \cdot 0,367 = 2,848 \cdot 10^{-11}$$

или, в логарифмических единицах,

$$(\sigma_{1\text{вых}}^{\text{acc}})^2, \text{ дБ} = 10 \lg(\sigma_{1\text{вых}}^{\text{acc}})^2 = -105,45 \text{ дБ}.$$

3. Анализ составляющих $e_{2\text{вых}}^{\text{acc}}(nT)$ и $e_{\text{вых}}^{\text{вых}}(nT)$.

Шумовые сигналы $e_{2\text{вых}}^{\text{acc}}(nT)$ и $e_{\text{вых}}^{\text{вых}}(nT)$ возникают непосредственно в точке выхода ЦПОС. Следовательно, справедливы равенства:

$$\begin{aligned} e_{2\text{вых}}^{\text{acc}}(nT) &= e_2^{\text{acc}}(nT) \quad \text{и} \quad e_{\text{вых}}^{\text{вых}}(nT) = e^{\text{вых}}(nT); \\ \mu_{2\text{вых}}^{\text{acc}} &= \mu_2^{\text{acc}} \quad \text{и} \quad \mu_{\text{вых}}^{\text{вых}} = \mu^{\text{вых}}; \\ (\sigma_{2\text{вых}}^{\text{acc}})^2 &= (\sigma_2^{\text{acc}})^2 \quad \text{и} \quad (\sigma_{\text{вых}}^{\text{вых}})^2 = (\sigma^{\text{вых}})^2. \end{aligned}$$

С учетом результатов проведенного анализа можно определить статистические характеристики суммарного шума квантования сигналов $e_{\text{вых}}(nT)$ в исследуемой системе:

- математическое ожидание сигнала $e_{\text{вых}}(nT)$

$$\mu_{\text{вых}} = \mu_{\text{вых}}^{\text{АПЧ}} + \mu_{1\text{вых}}^{\text{acc}} + \mu_{2\text{вых}}^{\text{acc}} + \mu_{\text{вых}}^{\text{вых}} = 0;$$

- дисперсия сигнала $e_{\text{вых}}(nT)$

$$\begin{aligned} (\sigma_{\text{вых}})^2 &= (\sigma_{\text{вых}}^{\text{АПЧ}})^2 + (\sigma_{1\text{вых}}^{\text{acc}})^2 + (\sigma_{2\text{вых}}^{\text{acc}})^2 + (\sigma_{\text{вых}}^{\text{вых}})^2 = \\ &= 2,848 \cdot 10^{-11} + 2,848 \cdot 10^{-11} + 7,761 \cdot 10^{-11} + 5,086 \cdot 10^{-6} \approx 5,086 \cdot 10^{-6} \end{aligned}$$

или, в логарифмических единицах,

$$(\sigma_{\text{вых}})^2, \text{ дБ} = 10 \lg(\sigma_{\text{вых}})^2 = -52,94 \text{ дБ}.$$

Полученные результаты полностью характеризуют эффекты квантования сигналов в исследуемой цифровой системе.

2.12. Адаптивные цифровые фильтры

Адаптивная обработка является универсальным средством решения проблемы априорной неопределенности и нестационарности обрабатываемого сигнала. Адаптивными называются системы, в которых получение желаемой реакции достигается путем преодоления недостатка априорной информации об обрабатываемом воздействии за счет более полного (по сравнению с неадаптивными системами) использования текущей информации. В системах инфокоммуникаций к таким системам можно отнести адаптивные антенные решетки, адаптивные компенсаторы сигналов электрического и акустического эха, адаптивные эквалайзеры (выравниватели) электрических характеристик каналов связи, адаптивные компенсаторы шумов.

Адаптацией называется регулярный процесс изменения параметров (или параметров и структуры) системы обработки, управляющих воздействий, осуществляемый на основе анализа текущей информации с целью достижения определенного (например оптимального) состояния системы. Процесс адаптации может быть реализован по входному сигналу $x(nT)$ или по выходному сигналу $y(nT)$. При решении задачи выбора метода адаптации учитываются свойства входного и выходного сигналов, тип параметров адаптации, сложность вычислений, устойчивость алгоритмов и т. д.

Системы, в которых процесс адаптации основан на использовании результатов анализа характеристик входного сигнала, называются адаптивными системами с управлением по входу.

Использование в адаптивных системах результатов анализа характеристик выходного сигнала переводит эти системы в класс адаптивных систем с управлением по выходу. Системы с управлением по выходу используются чаще, так как для них решены многие теоретические задачи, они удобны при проектировании нелинейных систем. Кроме того, такие системы хорошо приспособлены к работе в условиях сбоев, так как в них легко реализуются процедуры коррекции и переоптимизации. Адаптивным системам с управлением по выходу присущи и недостатки. В некоторых случаях рабочая функция, по которой реализуется адаптация системы с управлением по выходу, может иметь множество оптимумов, что затрудняет разработку алгоритмов адаптации. В других случаях в системе с управлением по выходу процесс адаптации может быть неустойчивым, что требует дополнительных ресурсов вычислителя для контроля устойчивости системы.

Одним из классов цифровых адаптивных систем являются адаптивные цифровые фильтры (АЦФ). Очевидно, что регулярное изменение параметров (параметров и структуры) адаптивного ЦФ требует наличия в таком фильтре устройства управления (устройства адаптации), в роли которого, как правило, выступают анализаторы входного или выходного сигналов, компараторы, минимизаторы ошибок и т. д. Обобщенная структурная схема адаптивного цифрового фильтра с управлением по выходу представлена на рис. 2.59.

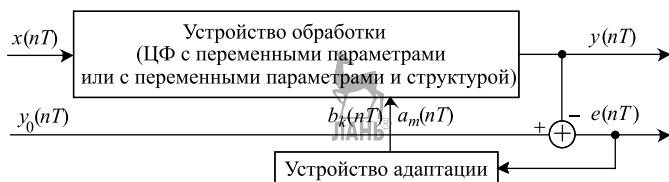


Рис. 2.59. Обобщенная структурная схема адаптивного цифрового фильтра с управлением по выходу

В адаптивном цифровом фильтре с управлением по выходу устройство обработки (цифровой фильтр с переменными параметрами или с переменными параметрами и структурой) предназначено для формирования из входного сигнала $x(nT)$ выходного сигнала $y(nT)$ таким образом, чтобы ошибка обработки $e(nT)$ была минимальной:

$$e(nT) = y_0(nT) - y(nT) \rightarrow \min_{b_k, a_m}, \quad (2.51)$$

где $y_0(nT)$ — эталонный (обучающий) сигнал.

Передаточная функция АЦФ определяется передаточной функцией устройства обработки: $H_{\text{АЦФ}}(z) = \frac{Y(z)}{X(z)}$.

Для получения сигнала $y_0(nT)$ используются различные приемы, целесообразность применения которых, как правило, определяется для каждой задачи проектирования индивидуально. Устройство адаптации обеспечивает перестройку параметров b_k и a_m (или параметров и структуры) устройства обработки на каждом интервале времени (одном или нескольких периодах дискретизации) таким образом, чтобы выполнялось условие (2.51). Естественно, что текущие характеристики АЦФ во временной, частотной и z -областях определяются значениями коэффициентов b_k и a_m , действующими на данном временном интервале.

Адаптивные ЦФ часто принадлежат к классу оптимальных систем. Следует иметь в виду, что оптимальными являются и некоторые системы, в которых параметры и структура устройства обработки неизменны. Они проектируются таким образом, чтобы их реакция $y(nT)$ была наилучшей в том или ином смысле среди всех возможных вариантов построения таких неизменяемых систем. Можно утверждать, что неперестраиваемые системы обработки сигналов являются оптимальными только при анализе реакции $y(nT)$ на бесконечно большом интервале времени. В адаптивной системе реакция $y(nT)$ на каждом интервале времени представляет собой наилучшую оценку желаемого эталонного сигнала $y_0(nT)$.

Адаптивные системы по своей природе являются изменяющимися во времени, как правило, нелинейными, их состояние зависит от изменения входного сигнала. Нелинейность оператора обработки в адаптивных системах имеет две особенности. Во-первых, в адаптивных системах нелинейность часто обусловлена процессом регулирования по характеристикам входного или выходного сигнала, полученным путем их усреднения на ограниченных интервалах времени. Во-вторых, нелинейность оператора обработки обусловлена целенаправленной перестройкой параметров и структуры устройства обработки, а иногда и устройства адаптации. Данная перестройка реализуется дискретно, что и порождает нелинейные свойства оператора обработки.

На рис. 2.60 в качестве примера показана структурная схема скалярного нерекурсивного параметрически адаптивного цифрового фильтра с управлением по выходу. Выходной сигнал такого фильтра можно определить в матричной форме:

$$y(nT) = \mathbf{X}^T(nT)\mathbf{H}(nT) = \mathbf{X}(nT)\mathbf{H}^T(nT),$$

где $\mathbf{X}(nT) = [x(nT), x(nT - T), x(nT - 2T), \dots, x(nT - KT + T)]$ — вектор сигналов, поступающих с выходов элементов задержки устройства обработки АЦФ; $\mathbf{H}(nT) = [b_0(nT), b_1(nT), b_2(nT), \dots, b_{K-1}(nT)]$ —

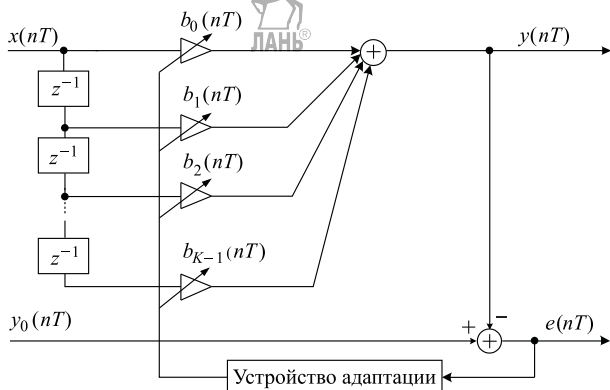


Рис. 2.60. Структурная схема нерекурсивного адаптивного цифрового фильтра с управлением по выходу

вектор отсчетов импульсной характеристики, равных (для НРЦФ) соответствующим весовым коэффициентам устройства обработки АЦФ.

Адаптивная цифровая фильтрация применяется в различных задачах обработки сигналов в системах телекоммуникаций. Примерами таких задач являются:

- цифровой спектральный анализ;
- компенсация искажений, вносимых при передаче полезного сигнала по каналу с замираниями;
- подавление отраженного сигнала (эхо-сигнала) в гибридной телефонной линии;
- анализ речевого сигнала в кодеках с линейным предсказанием.

Ниже кратко описаны два первых из перечисленных применений, при этом АЦФ структурно представлен в виде «черного ящика» со входами $x(nT)$ и $y_0(nT)$ и выходами $y(nT)$ и $e(nT)$ в соответствии с рис. 2.59 и 2.60. Особый интерес представляет тот факт, что применение АЦФ в рассматриваемых случаях построено на совершенно различных принципах и приводит к различным результатам.

1. Использование АЦФ для компенсации искажений, вносимых при передаче полезного сигнала по каналу с замираниями показано на рис. 2.61.

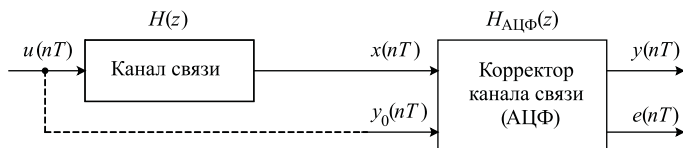


Рис. 2.61. Применение АЦФ для компенсации искажений

В общем случае передаточная функция $H(z)$ канала с замираниями не постоянна и флуктуирует под влиянием большого набора факторов. Сигнал, принимаемый из канала связи, поступает на вход $x(nT)$ АЦФ. Задачей устройства адаптации АЦФ является такое изменение передаточной функции $H_{\text{АЦФ}}(z)$, которое обеспечивало бы компенсацию канальных искажений и выполнение условия безыскаженной передачи сигналов $H_{\text{экв}}(z) = \frac{Y(z)}{U(z)} = H(z)H_{\text{АЦФ}}(z) = 1$. Очевидно, что данное условие будет выполняться при адаптивном формировании передаточной функции АЦФ в соответствии с требованием $H_{\text{АЦФ}}(z) = H^{-1}(z)$, которое и определяет закон функционирования АЦФ в рассмотренном случае. Так, если передающий тракт достаточно точно представляется передаточной функцией КИХ-фильтра, то устройство обработки компенсирующего АЦФ должно являться БИХ-фильтром с обратной передаточной функцией.

Сложность реализации рассматриваемого способа коррекции канала заключается в том, что совместное выполнение представленных выше условий предполагает следующие соотношения сигналов:

$$\begin{cases} y(nT) \approx y_0(nT); \\ y(nT) = u(nT), \end{cases}$$

что равносильно требованию $y_0(nT) \approx u(nT)$.

На практике обеспечение данного требования на всей длительности функционирования системы связи невозможно. Для решения указанного противоречия используется следующий приём: с временным интервалом, равным периоду адаптации, по каналу связи передаётся кратковременный тестовый (контрольный) сигнал. Знание его в точке приёма ($y_0(nT) = u(nT)$) позволяет осуществить адаптацию корректора, после чего канал связи вновь используется для передачи полезного сигнала.

Очевидно, что различные подходы к формированию АЦФ позволяют реализовать варианты амплитудной, фазовой и амплитудно-фазовой коррекции канала связи.

2. Использование АЦФ для компенсации телефонного эхо-сигнала показано на рис. 2.62. Сигнал, приходящий по приемной паре четырехпроводного канала связи, поступает на вход $x(nT)$ АЦФ. Гибридная мостовая схема, обеспечивающая согласование четырехпроводного тракта с двухпроводной абонентской линией, зачастую бывает разбалансированной, характеризуется передаточной функцией $H(z) = \frac{Y_0(z)}{X(z)}$ и является источником отраженного сигнала (эхо-сигнала), подаваемого на вход $y_0(nT)$ АЦФ.

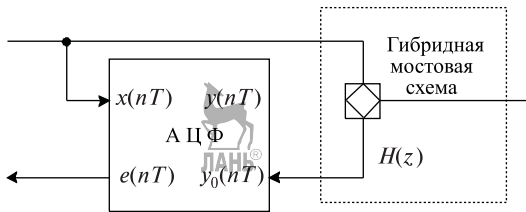


Рис. 2.62. Применение АЦФ для компенсации телефонного эхо-сигнала

Задачей устройства адаптации АЦФ является такое изменение $H_{\text{АЦФ}}(z)$, которое на каждом периоде адаптации обеспечивало бы максимальное совпадение сигналов $y_0(nT)$ и $y(nT)$. Полное равенство сигналов возможно при условии $H_{\text{АЦФ}}(z) = H(z)$, которое и определяет функционирование АЦФ в рассматриваемом случае. При соблюдении условия обеспечивается отсутствие эхо-сигнала на выходе $e(nT)$ АЦФ, являющемся в данном применении рабочим и соединенном с передающей парой четырехпроводного тракта.

Классификация адаптивных цифровых фильтров. Классификация адаптивных фильтров производится по различным классификационным признакам, в том числе по признакам, использованным ранее при рассмотрении систем классификации линейных дискретных систем и линейных цифровых фильтров. Ниже кратко перечислены некоторые, особенно часто используемые классы адаптивных ЦФ.

1. В зависимости от вида передаточных функций устройства обработки различают рекурсивные и нерекурсивные адаптивные цифровые фильтры.

2. В зависимости от способа реализации алгоритма адаптации выделяют адаптивные цифровые фильтры с управлением по входу и по выходу.

3. В зависимости от вида входного и выходного сигналов различают скалярные и векторные адаптивные цифровые фильтры. Если сигналы $x(nT)$ и $y(nT)$ — векторы, то фильтр считают векторным. Как правило, для описания векторных фильтров используется математический аппарат теории матриц.

4. В зависимости от степени адаптации различают параметрически адаптивные и структурно адаптивные ЦФ.

Необходимо отметить, что процесс адаптации может быть реализован путем вычисления рабочей функции адаптации как в частной, так и во временной областях. В соответствии с этим различают адаптивную фильтрацию в частотной и временной областях.

3 Цифровая многоскоростная обработка сигналов



3.1. Методы преобразования частоты. Классификация систем многоскоростной цифровой обработки сигналов

В современных инфокоммуникационных системах достаточно часто различные этапы обработки дискретных сигналов выполняются на разных частотах дискретизации f_d , что оказывается более эффективным, а в ряде случаев — необходимым. Такие системы получили название *систем с многочастотной дискретизацией* или *систем многоскоростной ЦОС*. Отличительным признаком таких систем является реализация процедуры преобразования частоты дискретизации, разновидностями которой являются *интерполяция* (повышение частоты дискретизации) и *децимация* (понижение частоты дискретизации). Соответствующие преобразования исходного дискретного сигнала $x(nT)$ с частотой дискретизации $f_d = 1/T$ в дискретный сигнал $y(n_y T_y)$ с частотой дискретизации $f_{dy} = 1/T_y$ осуществляются системами интерполяции и децимации, характеризующимися коэффициентами интерполяции $L = f_{dy}/f_d = T/T_y$ и децимации $M = f_d/f_{dy} = T_y/T$ (в простейшем случае — целочисленными).

Системы многоскоростной ЦОС, в которых увеличение (уменьшение) частоты дискретизации реализуется в один прием (однократно), называются *однократными*. Последовательное соединение однократных систем приводит к формированию *многократных* систем многоскоростной ЦОС.

Помимо систем интерполяции и децимации с целочисленными коэффициентами L и M существуют системы преобразования частоты дискретизации с рациональным коэффициентом L/M . Они формируются по принципу последовательного выполнения интерполяции и децимации, а следовательно, являются многократными многоскоростными системами ЦОС. Например, повышение частоты

дискретизации в 1,25 раза может быть реализовано последовательным соединением системы интерполяции с коэффициентом $L = 5$ и системы децимации с коэффициентом $M = 4$. В многократных системах при последовательном соединении систем интерполяции и децимации всегда первой ставится система интерполяции, а затем система децимации, что следует из анализа предназначения данных систем.

Следует учесть, что в теории цифровой обработки сигналов термин «интерполяция» не тождествен общепринятому математическому и не означает восстановление $L - 1$ отсчетов между соседними отсчетами исходного сигнала $x(nT)$, равно как термин «децимация» не означает отбрасывание $M - 1$ отсчетов исходного сигнала $x(nT)$ на периоде T_y . Эти термины в теории ЦОС используются в более общем смысле: рассматриваемые преобразования определяются как цифровая обработка входного сигнала $x(nT)$ с использованием операции линейной фильтрации, в результате которой формируется выходной сигнал $y(n_y T_y)$ с новой частотой дискретизации; при этом на временные и частотные изменения входного сигнала $x(nT)$ накладываются некоторые ограничения для обеспечения определенного соответствия характеристик выходного сигнала $y(n_y T_y)$ с соответствующими характеристиками сигнала $\tilde{y}(n_y T_y)$, который был бы получен путем непосредственной равномерной дискретизации исходного аналогового сигнала с частотой дискретизации $f_{ду}$ (рис. 3.1).

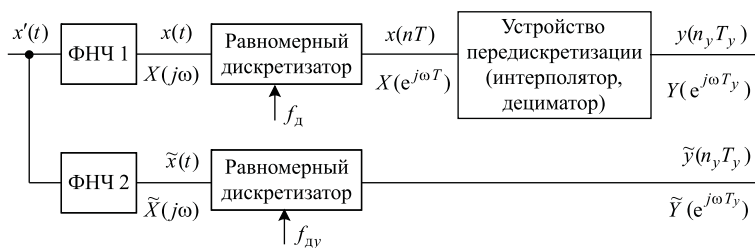


Рис. 3.1. Формирование дискретных сигналов процедурами передискретизации и непосредственной дискретизации

3.2. Однократные системы интерполяции и децимации

3.2.1. Однократная система интерполяции

Структурная схема однократной системы интерполяции с коэффициентом интерполяции L представлена на рис. 3.2.

Процедура интерполяции осуществляется в два этапа:

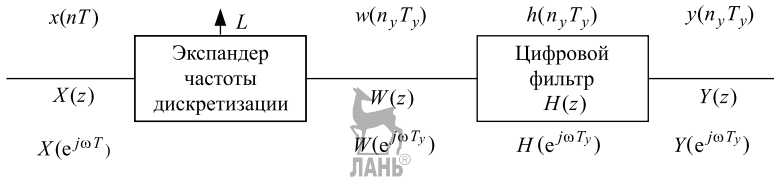


Рис. 3.2. Структурная схема однократной системы интерполяции

- на первом этапе из входного сигнала $x(nT)$ формируется промежуточный дискретный сигнал $w(n_y T_y)$ с частотой дискретизации $f_{\text{д}y} = L f_{\text{д}}$ и периодом дискретизации $T_y = T/L$;
- на втором этапе из промежуточного сигнала $w(n_y T_y)$ процедурой линейной цифровой фильтрации формируется требуемый выходной сигнал $y(n_y T_y)$.

Первый этап. Повышение частоты дискретизации обрабатываемого сигнала на первом этапе осуществляется экспандером частоты дискретизации, вставляющим $L - 1$ нулевых отсчетов, следующих с периодом T_y , между каждой парой отсчетов входного сигнала (рис. 3.3).

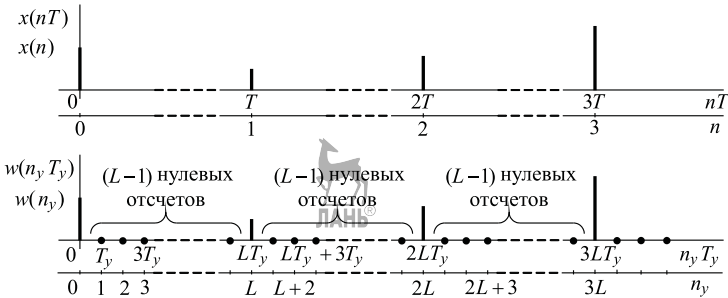


Рис. 3.3. Представление входного и промежуточного сигналов системы интерполяции во временной области

Соотношение нормированных временных шкал n и n_y определяется выражением

$$n = \frac{n_y}{L}, \quad n = 0, 1, 2, \dots; \quad n_y = 0, L, 2L, \dots \quad (3.1)$$

С учетом (3.1) выходной сигнал экспандера имеет вид

$$w(n_y) = \begin{cases} x(n_y/L), & n_y = 0, L, 2L, \dots; \\ 0, & n_y \neq 0, L, 2L, \dots \end{cases} \quad (3.2)$$

Определим z -изображение и спектр промежуточного сигнала $w(n_y)$. В соответствии с определением дискретного z -преобразования z -образы входного и промежуточного сигналов определяются

следующим образом:

$$X(z) = \sum_{n=0}^{\infty} x(n)z^{-n}; \quad W(z) = \sum_{n_y=0}^{\infty} w(n_y)z^{-n_y}.$$

Используя (3.2) и (3.1), можно преобразовать z -образ $W(z)$:

$$W(z) = \sum_{n_y=0, L, 2L, \dots}^{\infty} x\left(\frac{n_y}{L}\right) z^{-n_y} = \sum_{Ln=0, L, 2L, \dots}^{\infty} x(n) z^{-Ln} = \sum_{n=0}^{\infty} x(n) z^{-Ln}.$$

Следовательно, соотношение z -образов входного и выходного сигналов экспандера имеет вид

$$W(z) = X(z^L). \quad (3.3)$$

Выполнив замену $z = e^{j\omega T_y} = e^{j\omega_0 y}$, осуществим переход к спектрам исходного и промежуточного сигналов и получим их взаимосвязь:

$$W(e^{j\omega_0 y}) = X(e^{jL\omega_0 y}). \quad (3.4)$$

Известность соотношения периодов T и T_y позволяет определить зависимость между нормированными угловыми частотами $\omega_0 = \omega T$ и $\omega_{0y} = \omega T_y$ исходного и промежуточного сигналов соответственно: $\omega_0 = \omega T = \omega L T_y = L \omega_{0y}$. Подстановка этого выражения в (3.6) констатирует тождественность спектров сигналов $x(nT)$ и $w(n_y T_y)$:

$$W(e^{j\omega_{0y}}) = X(e^{j\omega_0}). \quad (3.5)$$

С учетом выражения (3.5) на рис. 3.4, *а, б, в* показаны следующие амплитудные спектры:

- $|X(j\omega)|$ — амплитудный спектр исходного аналогового сигнала $x(t)$, ограниченный частотой ω_{\max} ;
- $|X(e^{j\omega T})|$ — амплитудный спектр дискретного сигнала $x(nT)$, полученного из сигнала $x(t)$ процедурой дискретизации с частотой дискретизации f_d и поступающего на вход системы интерполяции;
- $|W(e^{j\omega T_y})|$ — амплитудный спектр промежуточного сигнала $w(n_y T_y)$, формирующегося на выходе экспандера.

Связь спектров входного и промежуточного сигналов интерполятора со спектром исходного аналогового сигнала определяется выражением

$$W(e^{j\omega T_y}) = X(e^{j\omega T}) = \frac{1}{T} \sum_{m=-\infty}^{\infty} X[j(\omega + m\omega_d)], \quad (3.6)$$

которое для основной полосы частот $\omega \in [0, \omega_d/2]$ имеет вид

$$W(e^{j\omega T_y}) = X(e^{j\omega T}) = \frac{1}{T} X(j\omega). \quad (3.7)$$

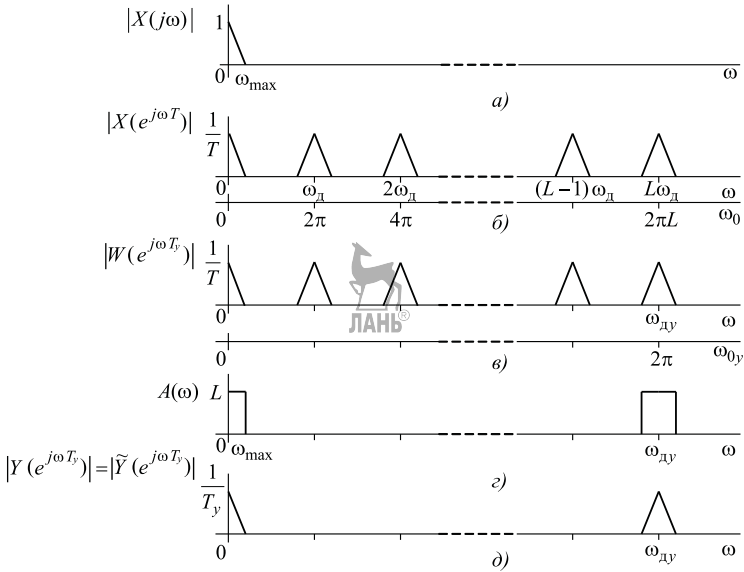


Рис. 3.4. Графическая иллюстрация процедуры интерполяции в частотной области для случая $\omega_d > 2\omega_{\max}$

Период спектра промежуточного сигнала $w(n_y T_y)$ определяется частотой дискретизации f_d входного сигнала.

Второй этап. На втором этапе интерполяции промежуточный сигнал $w(n_y T_y)$ подвергается процедуре линейной цифровой фильтрации с учетом следующего требования: амплитудный спектр $|Y(e^{j\omega T_y})|$ выходного сигнала интерполятора $y(n_y T_y)$ должен совпадать с амплитудным спектром $|\tilde{Y}(e^{j\omega T_y})|$ сигнала $\tilde{y}(n_y T_y)$, который был бы получен путем непосредственной дискретизации исходного аналогового сигнала $x(t)$ с частотой дискретизации $f_{дy}$. Связь спектра сигнала $\tilde{y}(n_y T_y)$ со спектром исходного аналогового сигнала определяется выражением

$$\tilde{Y}(e^{j\omega T_y}) = \frac{1}{T_y} \sum_{m=-\infty}^{\infty} X[j(\omega + m\omega_{дy})], \quad (3.8)$$

которое для основной полосы частот $\omega \in [0, \omega_{дy}/2]$ имеет вид

$$\tilde{Y}(e^{j\omega T_y}) = \frac{1}{T_y} X(j\omega). \quad (3.9)$$

Амплитудные спектры $|\tilde{Y}(e^{j\omega T_y})|$ и $|Y(e^{j\omega T_y})|$ показаны на рис. 3.4, д.

Сравнение амплитудных спектров $|W(e^{j\omega T_y})|$ промежуточного и $|Y(e^{j\omega T_y})|$ выходного сигналов интерполятора на периоде последнего

с учетом выражений (3.6) и (3.8) показывает, что их отличия состоят в следующем:

- амплитуды гармонических составляющих спектра $|W(e^{j\omega T_y})|$ масштабированы относительно соответствующих составляющих спектра $|X(j\omega)|$ коэффициентом $1/T = 1/LT_y$, а аналогичные составляющие спектра $|Y(e^{j\omega T_y})|$ — коэффициентом $1/T_y$;
- в наличии в спектре промежуточного сигнала $L - 1$ «лишних» спектральных компонентов шириной $2\omega_{\max}$.

Следовательно, для обеспечения равенства амплитудных спектров $|\tilde{Y}(e^{j\omega T_y})|$ и $|Y(e^{j\omega T_y})|$ задача линейной цифровой фильтрации второго этапа интерполяции состоит в подавлении «лишних» спектральных составляющих промежуточного сигнала и усилении его полезных спектральных составляющих в L раз. Такую задачу можно теоретически реализовать идеальным цифровым фильтром нижних частот (ФНЧ), амплитудно-частотная характеристика (АЧХ) которого (рис. 3.4,з) в основной полосе частот имеет вид

$$A(\omega) = \begin{cases} L, & 0 \leq \omega \leq \omega_{\max}; \\ 0, & \omega_{\max} < \omega < \omega_{dy}/2. \end{cases} \quad (3.10)$$

Очевидно, что в частотной области процедура интерполяции с учетом выражения (3.5) может быть описана следующим образом:

$$\begin{aligned} Y(e^{j\omega T_y}) &= W(e^{j\omega T_y})H(e^{j\omega T_y}) = X(e^{j\omega T})H(e^{j\omega T_y}) = \\ &= X(e^{j\omega T})|H(e^{j\omega T_y})|e^{j\arg\{H(e^{j\omega T_y})\}} = X(e^{j\omega T})A(\omega T_y)e^{j\varphi(\omega T_y)}, \end{aligned}$$

где $H(e^{j\omega T_y})$ — частотная характеристика ФНЧ;

$A(\omega T_y) = |H(e^{j\omega T_y})|$ — амплитудно-частотная характеристика ФНЧ;

$\varphi(\omega T_y) = \arg\{H(e^{j\omega T_y})\}$ — фазочастотная характеристика ФНЧ.

Следовательно, при идеальной АЧХ (3.10) в основной полосе частот соотношение спектров выходного и входного сигналов интерполятора определяется выражением

$$Y(e^{j\omega T_y}) = LX(e^{j\omega T})e^{j\varphi(\omega T_y)},$$

которое с учетом (3.7) принимает следующий вид:

$$Y(e^{j\omega T_y}) = \frac{L}{T}X(j\omega)e^{j\varphi(\omega T_y)} = \frac{1}{T_y}X(j\omega)e^{j\varphi(\omega T_y)}. \quad (3.11)$$

Сравнение (3.11) и (3.9) приводит к важному соотношению:

$$Y(e^{j\omega T_y}) = \tilde{Y}(e^{j\omega T_y})e^{j\varphi(\omega T_y)},$$

анализ которого позволяет сделать вывод, что спектр сигнала $y(n_y T_y)$ на выходе «идеальной» системы интерполяции можно рассматривать как спектр сигнала $\tilde{y}(n_y T_y)$, полученного путем непосредственной дискретизации исходного аналогового сигнала $x(t)$ с

частотой дискретизации $f_{\text{д}y}$, который изменен в соответствии с фазочастотной характеристикой (ФЧХ) ФНЧ интерполятора. Применение в интерполяторе ФНЧ с равномерной АЧХ и линейной ФЧХ обеспечивает «идеальную» форму сигнала $y(n_y T_y)$ и определяет реализацию в качестве ФНЧ нерекурсивного цифрового фильтра.

Следует учесть, что на практике фильтры, применяемые в интерполяторах, не имеют идеальной АЧХ и линейной ФЧХ, что обуславливает отличие и амплитудного, и фазового спектров сигнала $y(n_y T_y)$ от аналогичных спектров сигнала $\tilde{y}(n_y T_y)$.

Соотношение выходного и входного сигналов интерполятора во временной области нетрудно получить с использованием известной формулы линейной свертки:

$$\begin{aligned} y(n_y T_y) &= \sum_{m=0}^{\infty} w(m T_y) h(n_y T_y - m T_y) = \\ &= \sum_{m=0, 2L, \dots}^{\infty} x\left(\frac{m T_y}{L}\right) h(n_y T_y - m T_y). \end{aligned} \quad (3.12)$$

Очевидно, что в z -области соотношение выходного и входного сигналов интерполятора с учетом выражения (3.3) определяется следующим образом:

$$Y(z) = W(z)H(z) = X(z^L)H(z),$$

где $H(z)$ — передаточная функция цифрового фильтра.

3.2.2. Однократная система децимации

Структурная схема однократной системы децимации с целочисленным коэффициентом децимации M представлена на рис. 3.5.

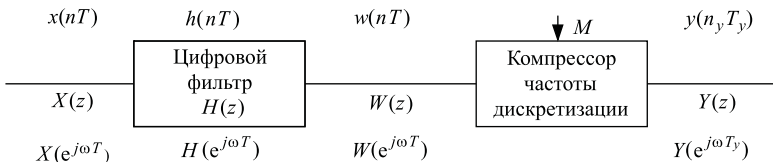


Рис. 3.5. Структурная схема однократной системы децимации

Процедура децимации осуществляется в два этапа:

- на первом этапе из входного сигнала $x(nT)$ процедурой линейной фильтрации формируется промежуточный дискретный сигнал $w(nT)$ с частотой дискретизации $f_{\text{д}} = 1/T$;
- на втором этапе из промежуточного сигнала $w(nT)$ формируется требуемый выходной сигнал $y(n_y T_y)$ с частотой дискретизации $f_{\text{д}y} = 1/T_y = 1/MT$.

Ограничение, накладываемое на решение задачи децимации средствами цифровой фильтрации, имеет следующий вид: амплитудный спектр $|Y(e^{j\omega T_y})|$ выходного сигнала дециматора $y(n_y T_y)$ должен совпадать с амплитудным спектром $|\tilde{Y}(e^{j\omega T_y})|$ сигнала $\tilde{y}(n_y T_y)$, который был бы получен путем непосредственной дискретизации исходного аналогового сигнала $x(t)$ с частотой дискретизации $f_{дy}$.

Первый этап. В частном случае, когда для формирования сигнала $x(nT)$ использовалась частота дискретизации $\omega_d = 2\pi f_d$ вдвое большая, чем верхняя частота ω_{\max} спектра исходного аналогового сигнала $x(t)$, спектр $X(e^{j\omega T})$ входного сигнала системы децимации занимает всю полосу частот и в основной полосе частот имеет верхнюю частоту $f_d/2$. Амплитудный спектр $|X(e^{j\omega T})|$ показан на рис. 3.6,а.

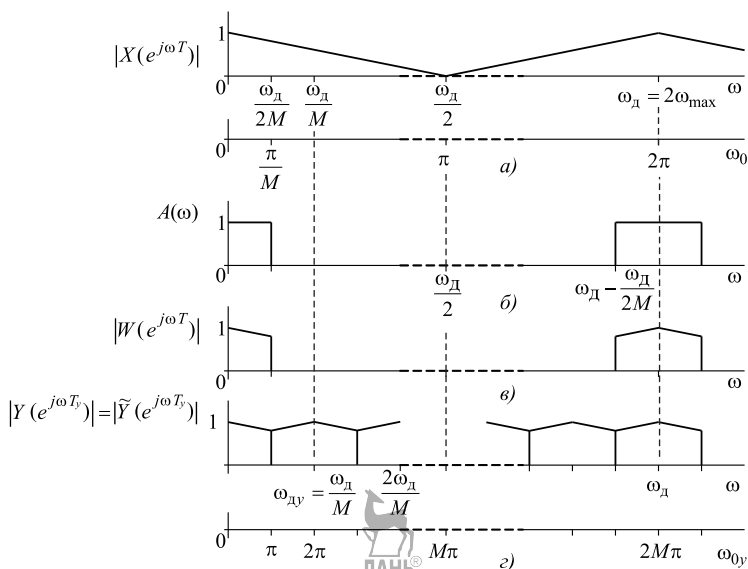


Рис. 3.6. Графическая иллюстрация процедуры децимации в частотной области для случая $\omega_d = 2\omega_{\max}$

Рассматриваемый случай является наиболее сложным для осуществления процедуры децимации, так как любое понижение частоты дискретизации повлечет за собой невыполнение требований теоремы Котельникова и обусловит явление наложения спектров. Исходя из этого, цифровая фильтрация первого этапа децимации нацелена на дополнительное ограничение спектра $X(e^{j\omega T})$ частотой $f_{дy}/2 = f_d/2M$, что равносильно дополнительному ограничению

спектра исходного аналогового сигнала такой частотой. Для этого в идеальном случае должен быть использован ФНЧ с амплитудно-частотной характеристикой (рис. 3.6,б)

$$A(\omega) = \begin{cases} 1, & 0 \leq \omega \leq \omega_d/2M; \\ 0, & \omega_d/2M \leq \omega \leq \omega_d/2. \end{cases} \quad (3.13)$$

Выполнение (3.13) гарантирует равенство амплитудных спектров промежуточного и входного сигналов дециматора в диапазоне $0 \leq \omega \leq \omega_d/2M$:

$$|W(e^{j\omega T})| = |X(e^{j\omega T})|. \quad (3.14)$$

Амплитудный спектр $|W(e^{j\omega T})|$ показан на рис. 3.6,в. Существо, что период этого спектра определяется частотой дискретизации ω_d исходного сигнала.

Во временной области формирование выходного сигнала ФНЧ представляется с помощью свертки входного сигнала с импульсной характеристикой фильтра:

$$w(nT) = \sum_{m=0}^{\infty} x(nT - mT)h(mT). \quad (3.15)$$

Второй этап. Формирование выходного сигнала $y(n_y T_y)$ из промежуточного сигнала $w(nT)$ производится в компрессоре частоты дискретизации (КЧД), осуществляющем процедуру прореживания сигнала $w(nT)$, в результате которой из последовательности входных отсчетов берется каждый M -й отсчет (рис. 3.7):

$$y(n_y T_y) = w(nT), \quad n = 0, M, 2M, \dots \quad (3.16)$$

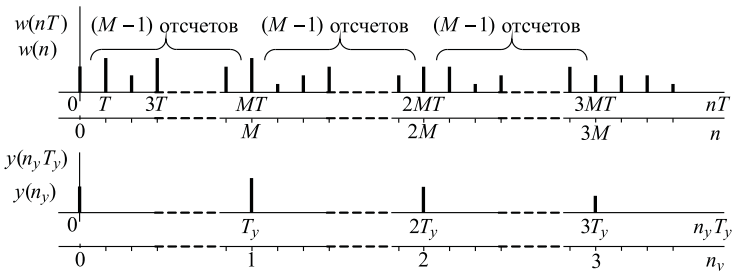


Рис. 3.7. Представление промежуточного и выходного сигналов системы децимации во временной области

Очевидно, что соотношение нормированных временных шкал n и n_y определяется выражением:

$$n = Mn_y, \quad n_y = 0, 1, 2, \dots; \quad n = 0, M, 2M, \dots \quad (3.17)$$

Объединение выражений (3.15), (3.16) и (3.17) позволяет получить соотношение входного и выходного сигналов системы децимации во временной области:

$$y(n_y T_y) = w(M n_y T) = \sum_{m=0}^{\infty} x(M n_y T - m T) h(m T); \quad n_y = 0, 1, 2, 3, \dots \quad (3.18)$$

Известно, что спектр выходного сигнала компрессора частоты дискретизации определяется масштабированной суммой спектров его входного сигнала, сдвинутых относительно друг друга по оси частот на величину $\omega_{\text{дy}}$:

$$Y(e^{j\omega T_y}) = \frac{1}{M} \sum_{m=0}^{M-1} W[e^{j(\omega + m\omega_{\text{дy}})T}].$$

Взаимосвязь спектров выходного $y(n_y T_y)$ и входного $x(nT)$ сигналов дециматора с учетом (3.14) и (3.18) определяется с помощью частотной характеристики $H(e^{j\omega T})$ цифрового фильтра дециматора следующим выражением:

$$Y(e^{j\omega T_y}) = \frac{1}{M} \sum_{m=0}^{M-1} X[e^{j(\omega + m\omega_{\text{дy}})T}] H[e^{j(\omega + m\omega_{\text{дy}})T}]. \quad (3.19)$$

Известность зависимости спектров аналогового сигнала и сигнала, полученного из него процедурой дискретизации, позволяет представить для основной полосы частот выражения, аналогичные (3.7) и (3.9):

$$X(e^{j\omega T}) = \frac{1}{T} X(j\omega); \quad (3.20)$$

$$\tilde{Y}(e^{j\omega T_y}) = \frac{1}{T_y} X(j\omega) = \frac{1}{MT} X(j\omega). \quad (3.21)$$

Очевидно, что при соблюдении ограничения (3.13) выражение (3.19) для основной полосы частот (при $m = 0$) с учетом выражения (3.20) упрощается до вида

$$Y(e^{j\omega T_y}) = \frac{1}{M} X(e^{j\omega T}) e^{j\varphi(\omega T)} = \frac{1}{MT} X(j\omega) e^{j\varphi(\omega T)}, \quad (3.22)$$

где $\varphi(\omega T) = \arg\{H(e^{j\omega T})\}$ — фазочастотная характеристика цифрового фильтра дециматора.

Сравнение формул (3.21) и (3.22) приводит к важному соотношению

$$Y(e^{j\omega T_y}) = \tilde{Y}(e^{j\omega T_y}) e^{j\varphi(\omega T)},$$

анализ которого позволяет сделать вывод, что спектр сигнала $y(n_y T_y)$ на выходе «идеальной» системы децимации можно рассматривать как спектр сигнала $\tilde{y}(n_y T_y)$, полученного путем непосредственной дискретизации исходного аналогового сигнала $x(t)$ с частотой дискретизации $f_{\text{ду}}$, который изменен в соответствии с фазочастотной характеристикой ФНЧ дециматора. Применение в дециматоре ФНЧ с равномерной АЧХ и линейной ФЧХ обеспечивает «идеальную» форму сигнала $y(n_y T_y)$, что определяет применение в качестве ФНЧ нерекурсивного цифрового фильтра.

Амплитудный спектр $|Y(e^{j\omega T_y})|$ выходного сигнала дециматора представлен на рис. 3.6,з. Он равен спектру амплитуд $|\tilde{Y}(e^{j\omega T_y})|$ сигнала $\tilde{y}(n_y T_y)$, который был бы получен путем непосредственной дискретизации исходного аналогового сигнала $x(t)$ с частотой дискретизации $f_{\text{ду}}$.

На практике фильтры, применяемые в интерполяторах, не имеют идеальной АЧХ и линейной ФЧХ, что обуславливает отличие и амплитудного, и фазового спектров сигнала $y(n_y T_y)$ от амплитудного и фазового спектров сигнала $\tilde{y}(n_y T_y)$.

3.3. Полифазные структуры систем интерполяции и децимации

Алгоритмы интерполяции и децимации, представленные в разд. 3.2.1 и 3.2.2, неэффективны в силу того, что содержат значительное количество ненужных вычислительных операций. Так, совершенно непроизводительной тратой вычислительных ресурсов на втором этапе процедуры интерполяции является обработка цифровым фильтром большого количества нулевых отсчетов промежуточного сигнала $w(n_y T_y)$. На втором этапе процедуры децимации происходит удаление части отсчетов промежуточного сигнала $w(nT)$, что делает абсолютно ненужными вычислительные операции, выполненные ранее цифровым фильтром для расчета удаляемых отсчетов. Поиск путей повышения эффективности организации вычислительных процедур изменения частоты дискретизации привел к разработке полифазных структур интерполяторов и дециматоров. Основная идея полифазных структур систем интерполяции и децимации состоит в замещении интерполяторов и дециматоров с одним цифровым нерекурсивным фильтром нижних частот (рис. 3.2, 3.5), работающим на высокой частоте дискретизации, эквивалентными системами с несколькими цифровыми ФНЧ с конечными импульсными характеристиками, работающими на низкой частоте дискретизации. Требуемое количество ФНЧ в полифазных структурах определяется коэффициентами интерполяции L и децимации M соответственно.

В основе функционирования полифазных структур интерполяторов и дециматоров лежит возможность разделения любого дискретного сигнала $x(n)$ на N составляющие (фазы) $x_k(n_1)$, $k = 0, 1, \dots, N - 1$. На рис. 3.8 показан пример разделения сигнала $x(n)$, $n = 0, 1, 2, \dots$, на три составляющие (при $N = 3$).

$$\begin{array}{cccccccccccccccc} x(n) & = & x(0), x(1), x(2), x(3), x(4), x(5), x(6), x(7), x(8), x(9), x(10), x(11), x(12), x(13), \dots \\ x_0(n_1) & = & x(0), & & x(3), & & x(6), & & x(9), & & x(12), \dots \\ x_1(n_1) & = & & x(1), & & x(4), & & x(7), & & x(10), & & x(13), \dots \\ x_2(n_1) & = & & & x(2), & & & x(5), & & & x(8), & & x(11), \dots \end{array}$$

Рис. 3.8. Разделение сигнала $x(n)$ на три составляющие (фазы)

В общем случае k -я составляющая $x_k(n_1)$ дискретного сигнала $x(n)$ имеет вид

$$x_k(n_1) = \{ \dots, x(k - N), x(k), x(k + N), \dots \} = \{ x(n_1 N + k) \},$$

а ее частота дискретизации в N раз меньше частоты дискретизации исходного сигнала $x(n)$. Z -образ составляющей $x_k(n_1)$ определяется выражением

$$X_k(z) = \sum_{n_1=0}^{\infty} x_k(n_1) z^{-n_1} = \sum_{n_1=0}^{\infty} x(n_1 N + k) z^{-n_1}.$$

На рис. 3.8 видно, что для того чтобы получить исходный сигнал $x(n)$ из отдельных фаз $x_k(n_1)$, необходимо вставить между отсчетами фаз по $N - 1$ нуля, задержать каждую фазу на число отсчетов, соответствующее ее номеру ($k = 0, 1, \dots, N - 1$), и суммировать задержанные последовательности. Из свойств z -преобразования известно, что вставка $N - 1$ нуля между отсчетами дискретного сигнала соответствует замене аргумента z -образа с z на z^N , а задержка дискретного сигнала на k периодов дискретизации обуславливает умножение z -образа на z^{-k} . Такие соответствия определяют связь z -преобразований исходного сигнала и отдельных его фаз:

$$X(z) = \sum_{k=0}^{N-1} X_k(z^N) z^{-k}.$$

Структурная схема системы, разделяющей дискретный сигнал на N составляющих, а затем формирующей его из отдельных фаз, показана на рис. 3.9. Следует отметить, что последовательное осуществление указанных операций приводит к суммарной задержке выходного сигнала системы относительно входного на $N - 1$ периода дискретизации.

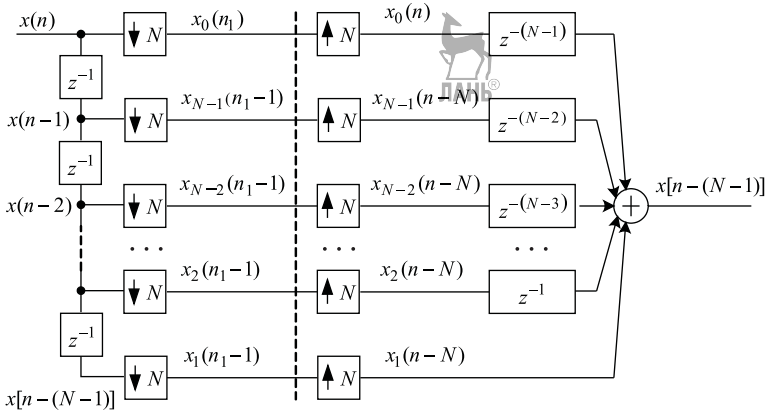


Рис. 3.9. Структурная схема системы, разделяющей дискретный сигнал на фазы (слева) и формирующей его из отдельных фаз (справа)

Полифазная структура систем интерполяции. Для демонстрации перехода к полифазным структурам рассмотрим функционирование однократной системы интерполяции с коэффициентом $L = 2$ (рис. 3.10), содержащей нерекурсивный цифровой ФНЧ 5-го порядка. Отсчеты конечной импульсной характеристики такого фильтра определяются его коэффициентами: $h(n_y T_y) = (b_0, b_1, b_2, b_3, b_4, b_5)$.

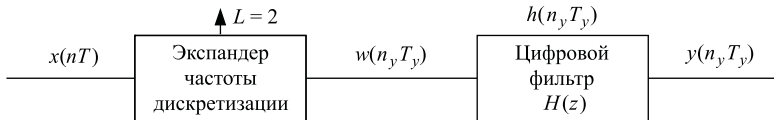


Рис. 3.10. Система интерполяции с коэффициентом $L = 2$

Пусть входной сигнал системы интерполяции имеет вид $x(n) = (1, 3, 2, -4, 1, 3, -1, \dots)$, тогда на выходе экспандера частоты дискретизации в соответствии с выражением (3.2) формируется сигнал $w(n_y T_y) = (1, 0, 3, 0, 2, 0, -4, 0, 1, 0, 3, 0, -1, 0, \dots)$.

Выходной сигнал рассматриваемой системы интерполяции формируется процедурой цифровой фильтрации на основании формулы (3.12) и представляет собой свертку промежуточного сигнала и импульсной характеристики фильтра:

$$y(n_y T_y) = \sum_{m=0}^{n_y} w(m T_y) h(n_y T_y - m T_y).$$

Рассмотрим детально процедуру формирования отсчетов выходного сигнала, подчеркивая отсчеты промежуточного сигнала, име-

ющие нулевые значения:

$$\begin{aligned} y(0T_y) &= w(0T_y)h(0T_y - 0T_y) = w(0)h(0) = w(0)b_0; \\ y(1T_y) &= w(0T_y)h(1T_y - 0T_y) + w(1T_y)h(1T_y - 1T_y) = \\ &= w(0)b_1 + \underline{w(1)b_0}; \\ y(2T_y) &= w(0T_y)h(2T_y - 0T_y) + w(1T_y)h(2T_y - 1T_y) + \\ &+ w(2T_y)h(2T_y - 2T_y) = w(0)b_2 + \underline{w(1)b_1} + w(2)b_0. \end{aligned}$$

Аналогично представим выражения нескольких последующих отсчетов:

$$\begin{aligned} y(3T_y) &= w(0)b_3 + \underline{w(1)b_2} + w(2)b_1 + \underline{w(3)b_0}; \\ y(4T_y) &= w(0)b_4 + \underline{w(1)b_3} + w(2)b_2 + \underline{w(3)b_1} + w(4)b_0; \\ y(5T_y) &= w(0)b_5 + \underline{w(1)b_4} + w(2)b_3 + \underline{w(3)b_2} + w(4)b_1 + \underline{w(5)b_0}; \\ y(6T_y) &= \underline{w(1)b_5} + \underline{w(2)b_4} + \underline{w(3)b_3} + w(4)b_2 + \underline{w(5)b_1} + w(6)b_0; \\ y(7T_y) &= w(2)b_5 + \underline{w(3)b_4} + w(4)b_3 + \underline{w(5)b_2} + w(6)b_1 + \underline{w(7)b_0}; \dots \end{aligned}$$

С учетом известного соотношения сигналов $x(n)$ и $w(n_y T_y)$ для ненулевых значений промежуточного сигнала можно записать: $w(0) = x(0)$, $w(2) = x(1)$, $w(4) = x(2)$, $w(6) = x(3)$ и т. д. Тогда отсчеты выходного сигнала определяются следующим образом:

$$\begin{aligned} y(0T_y) &= x(0)b_0; \\ y(1T_y) &= x(0)b_1; \\ y(2T_y) &= x(0)b_2 + x(1)b_0; \\ y(3T_y) &= x(0)b_3 + x(1)b_1; \\ y(4T_y) &= x(0)b_4 + x(1)b_2 + x(2)b_0; \\ y(5T_y) &= x(0)b_5 + x(1)b_3 + x(2)b_1; \\ y(6T_y) &= x(1)b_4 + x(2)b_2 + x(3)b_0; \\ y(7T_y) &= x(1)b_5 + x(2)b_3 + x(3)b_1; \dots \end{aligned}$$

Особенность представленных выражений в том, что они показывают возможность расчета выходного сигнала интерполятора лишь на основе знания входного сигнала интерполятора и импульсной характеристики фильтра (в рассматриваемом примере — коэффициентов фильтра), что устраняет необходимость осуществления операций экспандирования и последующей фильтрации нулевых отсчетов сигнала $w(n_y T_y)$.

Анализ представленных выражений свидетельствует о том, что в нерекурсивном цифровом фильтре 5-го порядка для формирования четных отсчетов выходного сигнала используются коэффициенты b_0, b_2, b_4 , а для формирования нечетных отсчетов — коэффициен-

ты b_1, b_3, b_5 , масштабирующие ненулевые отсчеты сигнала $w(n_y T_y)$. Следовательно, возникает возможность разделить (заменить) исходный НРЦФ 5-го порядка, работающий на выходной (высокой) частоте дискретизации, на два НРЦФ 3-го порядка с импульсными характеристиками $h_1(nT) = (b_0, b_2, b_4)$ и $h_2(nT) = (b_1, b_3, b_5)$. Очевидно, что данные фильтры будут функционировать на входной (низкой) частоте дискретизации (рис. 3.11), а на их выходах будут формироваться дискретные сигналы $y_1(nT)$ и $y_2(nT)$, представляющие собой составляющие (фазы) выходного сигнала:

$$y_1(nT) = y(0T_y), y(2T_y), y(4T_y), y(6T_y), \dots$$

$$y_2(nT) = y(1T_y), y(3T_y), y(5T_y), y(7T_y), \dots$$

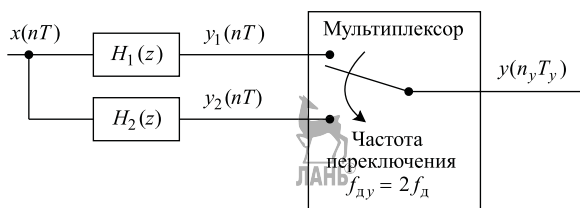


Рис. 3.11. Полифазная структурная схема интерполятора с коэффициентом $L = 2$

Для формирования выходного сигнала $y(n_y T_y)$ с высокой частотой дискретизации в этом случае необходимо поочередно считывать отсчеты с выходов обоих фильтров, при этом частота опроса фильтров мультиплексором должна вдвое превышать частоту функционирования фильтров (частоту дискретизации входного сигнала интерполятора).

Очевидно, что процедура интерполяции на основе полифазной структуры характеризуется меньшей вычислительной сложностью по сравнению со структурой, представленной на рис. 3.10.

Ниже представлено обоснование возможности перехода к полифазным структурам интерполяции в общем виде.

Прямое использование выражения (3.12) позволяет рассчитать значения отсчетов интерполированного дискретного сигнала $y(n_y)$ при нулевых начальных условиях цифрового фильтра (при расчетах учтено, что при $n_y - m \leq 0$ значения импульсной характеристики $h(n_y - m) = 0$):

- для $n_y = 0, 1, 2, \dots, L - 1$:

$$y(0) = h(0)x(0);$$

$$y(1) = h(1)x(0);$$

$$\dots\dots\dots$$

$$y(L - 1) = h(L - 1)x(0);$$

- для $n_y = L, L + 1, L + 2, \dots, 2L - 1$:

$$y(L) = h(L)x(0) + h(0)x(1);$$

$$y(L + 1) = h(L + 1)x(0) + h(1)x(1);$$

.....

$$y(2L - 1) = h(2L - 1)x(0) + h(L - 1)x(1);$$

- для $n_y = 2L, 2L + 1, 2L + 2, \dots, 3L - 1$:

$$y(2L) = h(2L)x(0) + h(L)x(1) + h(0)x(2);$$

$$y(2L + 1) = h(2L + 1)x(0) + h(L + 1)x(1) + h(1)x(2);$$

.....

$$y(3L - 1) = h(3L - 1)x(0) + h(2L - 1)x(1) + h(L - 1)x(2)$$

и т. д.

Применение в качестве ФНЧ нерекурсивного цифрового фильтра приводит к тому, что отсчеты импульсной характеристики равны соответствующим коэффициентам фильтра, а максимальное количество слагаемых при расчете отсчетов сигнала $y(n_y)$ будет зависеть от порядка фильтра (будет определяться количеством отсчетов конечной импульсной характеристики).

Использование идеи полифазного представления дискретных сигналов делает возможным разделение сигнала $y(n_y)$ на L фаз $y_l(n)$, $l = 0, 1, 2, \dots, L - 1$, с частотой дискретизации f_d (с частотой дискретизации исходного сигнала $x(nT)$):

$$y_0(n) = y(0), y(L), y(2L), \dots;$$

$$y_1(n) = y(1), y(L + 1), y(2L + 1), \dots;$$

$$y_2(n) = y(2), y(L + 2), y(2L + 2), \dots;$$

.....

$$y_l(n) = y(l), y(L + l), y(2L + l), \dots;$$

.....

$$y_{L-1}(n) = y(L - 1), y(2L - 1), y(3L - 1), \dots$$

Например, отсчеты составляющей $y_0(n)$ определяются следующим образом:

$$\begin{cases} y_0(0) = y(0) = h(0)x(0); \\ y_0(1) = y(L) = h(L)x(0) + h(0)x(1); \\ y_0(2) = y(2L) = h(2L)x(0) + h(L)x(1) + h(0)x(2); \\ \dots \end{cases} \quad (3.23)$$

Такой подход позволяет сформировать сигналы $y_l(n)$ на выходах L цифровых ФНЧ (рис. 3.12), работающих на низкой частоте дискретизации, на вход которых поступает исходный сигнал $x(n)$ системы интерполяции. В этом случае правомерно говорить о замене в

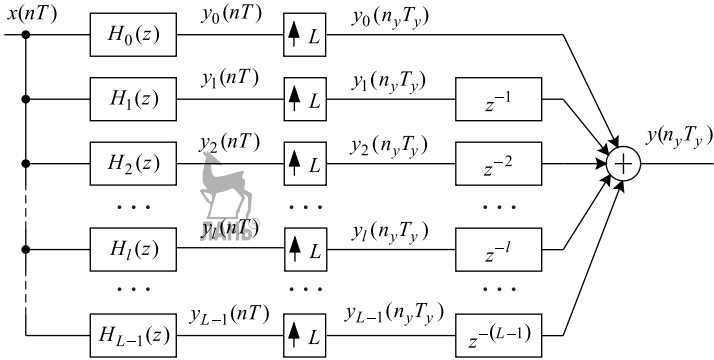


Рис. 3.12. Полифазная структура системы интерполяции

системе интерполяции с коэффициентом интерполяции L цифрового КИХ-фильтра порядка $K - 1$ с коэффициентами $\{h(0), h(1), h(2), \dots, h(L), \dots, h(K-1)\}$ на L цифровых КИХ-субфильтрах порядка $K/L - 1$ (очевидно, что отношение K/L должно быть целым числом). При обозначении порядкового номера субфильтра как $l = 0, 1, 2, \dots, L-1$ и известности выходного сигнала субфильтра легко определяется его импульсная характеристика, а следовательно, и его коэффициенты. Так, на основании (3.23) импульсная характеристика нулевого субфильтра (при $l = 0$) имеет вид

$$h_0(n) = \{h(0), h(L), h(2L), \dots, h[(K/L - 1)L]\}.$$

Аналогично определяются импульсные характеристики остальных субфильтров:

$$h_1(n) = \{h(1), h(L+1), h(2L+1), \dots, h[(K/L - 1)L + 1]\};$$

$$h_2(n) = \{h(2), h(L+2), h(2L+2), \dots, h[(K/L - 1)L + 2]\};$$

.....

$$h_l(n) = \{h(l), h(L+l), h(2L+l), \dots, h[(K/L - 1)L + l]\};$$

.....

$$h_{L-1}(n) = \{h(L-1), h(2L-1), h(3L-1), \dots, h[(K/L - 1)L + (L-1)]\}.$$

Z -преобразования импульсных характеристик позволяют определить передаточные функции субфильтров:

$$H_0(z) = h(0) + h(L)z^{-1} + h(2L)z^{-2} + \dots + h[(K/L - 1)L]z^{-(K/L-1)};$$

$$H_1(z) = h(1) + h(L+1)z^{-1} + \dots + h[(K/L - 1)L + 1]z^{-(K/L-1)};$$

.....

$$H_l(z) = h(l) + h(L+l)z^{-1} + \dots + h[(K/L - 1)L + l]z^{-(K/L-1)};$$

.....

$$H_{L-1}(z) = h(L-1) + h(2L-1)z^{-1} + \dots + h[(K/L-1)L + (L-1)]z^{-(K/L-1)}.$$

На рис. 3.12 представлена полифазная структура системы интерполяции, состоящая из набора субфильтров, экспандеров частоты дискретизации, линий задержки и сумматора.

Существенно, что частота дискретизации выходных сигналов $y_l(n)$ цифровых субфильтров равна частоте дискретизации f_d исходного сигнала $x(nT)$. Повышение этой частоты до требуемой f_{dy} осуществляется экспандерами частоты дискретизации с коэффициентами L , включенными последовательно с цифровыми фильтрами. Таким образом, в интерполяторах с полифазной структурой вставки нулей в ходе процедуры повышения частоты дискретизации осуществляются уже после цифровой фильтрации, что обеспечивает устранение ненужных операций обработки цифровым фильтром нулевых отсчетов промежуточного сигнала $w(n_y T_y)$.

Выходной сигнал $y(n_y T_y)$ системы интерполяции формируется в виде суммы выходных сигналов $y_l(n_y T_y)$ экспандеров, сдвинутых друг относительно друга на период T_y , что обеспечивается $L-1$ линиями задержки на $1, 2, \dots, L-1$ периодов T_y .

Анализ процедур обработки сигнала в интерполяторе с полифазной структурой (см. рис. 3.12) с учетом свойств z -преобразования позволяет представить процесс интерполяции в z -области:

- z -образы сигналов $y_l(nT)$ на выходах субфильтров определяются как $Y_l(z) = X(z)H_l(z)$;
- z -образы сигналов $y_l(n_y T_y)$ на выходах экспандеров определяются как $Y'_l(z) = Y_l(z^L) = X(z^L)H_l(z^L)$;
- z -образ выходного сигнала $y(n_y T_y)$ на выходе интерполятора с полифазной структурой определяется следующим выражением:

$$Y(z) = \sum_{l=0}^{L-1} Y'_l(z)z^{-l} = \sum_{l=0}^{L-1} X(z^L)H_l(z^L)z^{-l}. \quad (3.24)$$

Выражение (3.24) определяет для интерполятора с полифазной структурой связь выходного и входного сигналов в z -области.

Очевидно, что операции сложения сигналов $y_0(n_y T_y), y_1(n_y T_y - T_y), \dots, y_l(n_y T_y - lT_y), \dots, y_{L-1}[n_y T_y - (L-1)T_y]$ являются избыточными, так как при вычислении каждого отсчета выходного сигнала $y(n_y T_y)$ ненулевое значение присутствует только на одном из входов сумматора. По этой причине на практике чаще используется коммутационная полифазная структура системы интерполяции, представленная на рис. 3.13.

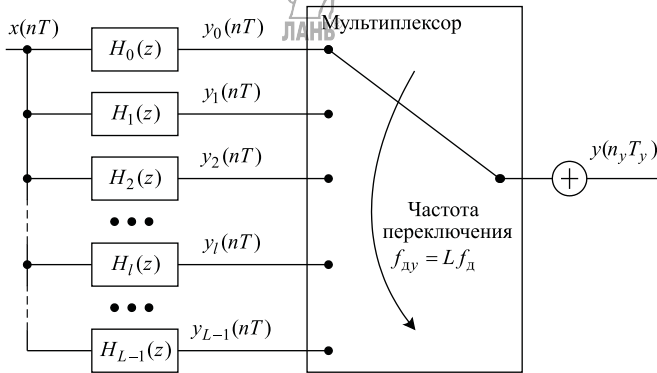


Рис. 3.13. Коммутационная полифазная структура системы интерполяции

При такой структуре экспандеры частоты дискретизации, линии задержки и сумматор заменяются, по сути дела, процедурой мультиплексирования: на выходах субфильтров *одновременно* формируются L отсчетов интерполированного сигнала, которые за период T поочередно подаются на выход, что и обеспечивает повышение частоты дискретизации до $f_{дл} = L f_д$. В этом случае очевидным является требование сохранения сигналов $y_l(nT)$ на выходах субфильтров до момента опроса l -го подканала (до момента считывания l -й фазы).

Полифазная структура систем децимации. В п. 3.2.2 показано, что соотношение выходного и входного сигналов дециматора с коэффициентом децимации M (см. рис. 3.5) определяется выражением (3.18), применение которого позволяет рассчитать значения отсчетов сигнала $y(n_y)$ на выходе системы децимации при нулевых начальных условиях цифрового фильтра нижних частот:

$$y(0) = h(0)x(0);$$

$$y(1) = h(0)x(M) + h(1)x(M-1) + h(2)x(M-2) + \dots + h(M)x(0);$$

$$y(2) = h(0)x(2M) + h(1)x(2M-1) + h(2)x(2M-2) + \dots + h(2M)x(0);$$

$$y(3) = h(0)x(3M) + h(1)x(3M-1) + h(2)x(3M-2) + \dots + h(3M)x(0);$$

.....

$$y(M) = h(0)x(M^2) + h(1)x(M^2-1) + h(2)x(M^2-2) + \dots + h(M^2)x(0);$$

.....

В общем случае справедливо выражение

$$y(n_y) = h(0)x(n_y M) + h(1)x(n_y M - 1) + h(2)x(n_y M - 2) + \dots + h(n_y M)x(0). \quad (3.25)$$

Применение в качестве ФНЧ цифрового нерекурсивного филь-

тра приводит к тому, что отсчеты импульсной характеристики равны соответствующим коэффициентам фильтра, а максимальное количество слагаемых при расчете отсчетов сигнала $y(n_y)$ по формуле (3.25) будет зависеть от порядка $K - 1$ фильтра (будет определяться количеством отсчетов конечной импульсной характеристики).

Анализ выражения (3.25) показывает возможность получения сигнала $y(n_y)$ в виде суммы сигналов $y_m(n_y)$, сформированных на выходах M цифровых субфильтров, функционирующих на частоте дискретизации $f_{\text{дц}}$ сигнала $y(n_y)$ и имеющих порядок $K/M - 1$ (при обязательном равенстве отношения K/M целому числу). Математическим пояснением такой возможности является представление выражения (3.25) в виде суммы M слагаемых:

$$y(n_y) = y_0(n_y) + y_1(n_y) + y_2(n_y) + \dots + y_m(n_y) + \dots + y_{M-1}(n_y),$$

где

$$\begin{aligned} y_0(n_y) = & h(0)x(n_y M) + h(M)x(n_y M - M) + h(2M)x(n_y M - 2M) + \\ & + h(3M)x(n_y M - 3M) + \dots + h[(K/M - 1)M] \times \\ & \times x[n_y M - (K/M - 1)M]; \end{aligned}$$

$$\begin{aligned} y_1(n_y) = & h(1)x(n_y M - 1) + h(1 + M)x[n_y M - (1 + M)] + \\ & + h(1 + 2M)x[n_y M - (1 + 2M)] + \dots + h[1 + (K/M - 1)M] \times \\ & \times x[n_y M - [1 + (K/M - 1)M]]; \end{aligned}$$

.....

$$\begin{aligned} y_m(n_y) = & h(m)x(n_y M - m) + h(m + M)x[n_y M - (m + M)] + \\ & + h(m + 2M)x[n_y M - (m + 2M)] + \dots + h[m + (K/M - 1)M] \times \\ & \times x[n_y M - [m + (K/M - 1)M]]; \end{aligned}$$

.....

$$\begin{aligned} y_{M-1}(n_y) = & h(M - 1)x[n_y M - (M - 1)] + h(2M - 1) \times \\ & \times x[n_y M - (2M - 1)] + h(3M - 1)x[n_y M - (3M - 1)] + \dots + \\ & + h(K - 1)x[n_y M - (K - 1)]. \end{aligned}$$

Тогда очевидно, что субфильтры с порядковыми номерами $m = 0, 1, 2, \dots, M - 1$ должны иметь конечные импульсные характеристики $h_m(n_y)$, отсчеты которых определяются соответствующими отсчетами импульсной характеристики $h(n)$ исходного цифрового КИХ-фильтра (см. рис. 3.5):

$$h_0(n_y) = \{h(0), h(M), h(2M), \dots, h[(K/M - 1)M]\};$$

$$h_1(n_y) = \{h(1), h(1 + M), h(1 + 2M), \dots, h[1 + (K/M - 1)M]\};$$

$$h_2(n_y) = \{h(2), h(2 + M), h(2 + 2M), \dots, h[2 + (K/M - 1)M]\};$$

.....

$$h_m(n_y) = \{h(m), h(m+M), h(m+2M), \dots, h[m+(K/M-1)M]\};$$

.....

$$h_{M-1}(n_y) = \{h(M-1), h[(M-1)+M], h[(M-1)+2M], \dots, h(K-1)\},$$

а в качестве входных сигналов субфильтров должны выступать дискретные последовательности $x_m(n_y)$, отсчеты которых определяются отсчетами входного сигнала $x(n)$ системы децимации:

$$\begin{cases} x_0(n_y) = \{x(0), x(M), x(2M), x(3M), \dots\}; \\ x_1(n_y) = \{0, x(M-1), x(2M-1), x(3M-1), \dots\}; \\ \dots\dots\dots \\ x_m(n_y) = \{0, x(M-m), x(2M-m), x(3M-m), \dots\}; \\ \dots\dots\dots \\ x_{M-1}(n_y) = \{0, x(1), x(M+1), x(2M+1), \dots\}. \end{cases}$$

Передаточные функции $H_m(z)$ субфильтров определяются z -преобразованием конечных импульсных характеристик $h_m(n_y)$:

$$\begin{aligned} H_0(z) &= h(0) + h(M)z^{-1} + h(2M)z^{-2} + \dots + h[(K/M-1)M] \times \\ &\quad \times z^{-(K/M-1)}; \\ H_1(z) &= h(1) + h(1+M)z^{-1} + \dots + h[1+(K/M-1)M]z^{-(K/M-1)}; \\ &\dots\dots\dots \\ H_m(z) &= h(m) + h(m+M)z^{-1} + \dots + h[m+(K/M-1)M]z^{-(K/M-1)}; \\ &\dots\dots\dots \\ H_{M-1}(z) &= h(M-1) + h(2M-1)z^{-1} + \dots + h(K-1)z^{-(K/M-1)}. \end{aligned} \quad (3.26)$$

Анализ дискретных сигналов $x_m(n_y)$, определяемых по выражению (3.26), позволяет сделать вывод о том, что они в соответствии с полифазным представлением сигналов представляют собой совокупность M составляющих (фаз) сигнала $x(n)$, полученных системой разделения (см. левую часть рис. 3.9).

Изложенный материал позволяет представить полифазную структуру системы децимации (рис. 3.14).

Цифровые субфильтры, входящие в состав полифазной структуры, функционируют на низкой частоте дискретизации $f_{\text{ду}}$. Отбрасывание лишних отсчетов сигнала $x(n)$ в представленной структуре производится компрессорами частоты дискретизации с коэффициентами M до процедуры фильтрации, что позволяет избавиться от ненужных операций по их расчету, выполняемых цифровым фильтром дециматора, имеющего структуру, показанную на рис. 3.5.

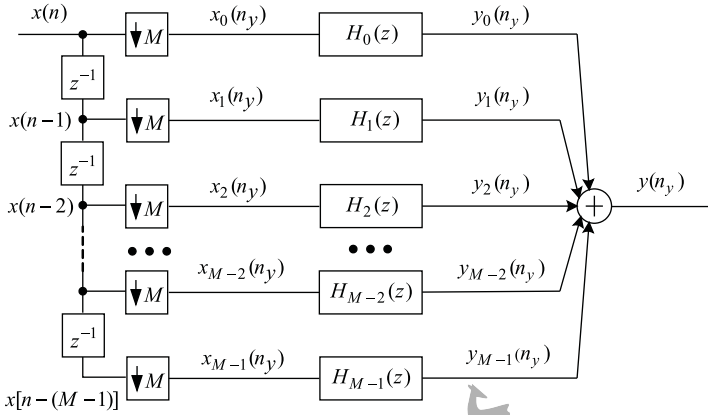


Рис. 3.14. Полифазная структура системы децимации

Назначение линий задержки и компрессоров частоты дискретизации в схеме на рис. 3.14 состоит в одновременной подаче входных сигналов $x_m(n_y)$ на цифровые субфильтры. На практике обычно используется коммутационная полифазная структура системы децимации (рис. 3.15), в которой эту роль выполняет демультиплексор, осуществляя необходимую последовательность коммутации отсчетов входного сигнала $x(n)$ на входы субфильтров. Естественным требованием в этом случае является сохранение сигналов $y_m(n_y)$ на выходах субфильтров в течение полного цикла коммутации (периода дискретизации T_y) для обеспечения корректного формирования выходного сигнала $y(n_y)$.

В качестве примера рассмотрим возможность построения полифазной структуры системы децимации на основе однократной системы с коэффициентом $M = 2$. Низкочастотный фильтр в однократной системе выполнен на базе нерекурсивного КИХ-фильтра 5-го порядка. Импульсная характеристика такого фильтра $h(n) = (b_0, b_1, b_2, b_3, b_4, b_5)$ определяется значениями его коэффициентов.

В качестве входного сигнала на вход системы подается дискретный сигнал $x(n)$, тогда при осуществлении цифровой фильтрации свертка импульсной характеристики фильтра и входного сигнала выглядит следующим образом:

$$w(nT) = \sum_{m=0}^n x(nT - mT)h(mT). \quad (3.27)$$

В компрессоре частоты дискретизации осуществляется обнуление $M - 1$ промежуточных отсчетов (для данного примера каждого четного отсчета) и, соответственно, переход к более низкой частоте

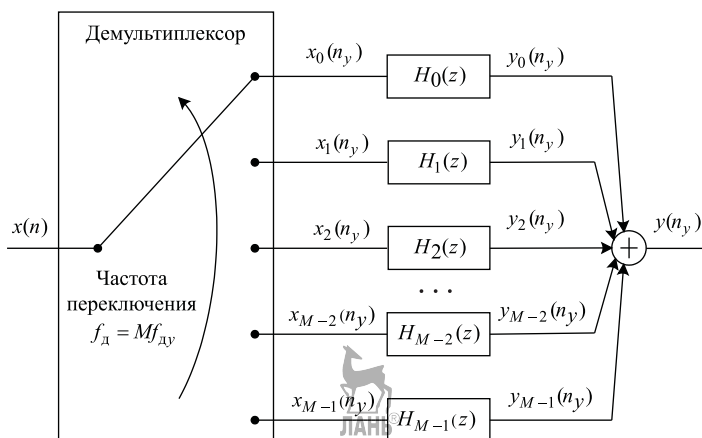


Рис. 3.15. Коммутационная полифазная структура системы децимации

дискретизации (на основе учета отсчетов, не подвергаемых обнулению). Таким образом, в НРЦФ 5-го порядка рассчитываются отсчеты, которые в дальнейшем не используются для формирования выходного сигнала. Основная идея перехода к полифазным системам децимации — осуществлять расчет «нужных» отсчетов сигнала, которые и будут формировать выходной сигнал. Для этого можно рассмотреть выражение (3.27). Данная операция проводится на высокой (входной) частоте дискретизации, но формирование «нужного» отсчета будет происходить в M раз медленнее, чем поступление новых отсчетов входного сигнала.

Таким образом, для того чтобы сформировать только «нужный» отсчет выходного сигнала с использованием процедуры фильтрации на низкой (выходной) частоте дискретизации, необходимо разделить исходную ИХ НРЦФ на два нерекурсивных цифровых субфильтра с коэффициентами $h_1(nT) = (b_0, b_2, b_4)$ и $h_2(nT) = (b_1, b_3, b_5)$. В соответствии с анализом формулы (3.27) на данные фильтры необходимо подавать попеременно отсчеты входного сигнала с частотой входного сигнала системы децимации, а выходы цифровых фильтров, работающих на частоте дискретизации выходного сигнала системы децимации, объединять, что и показано на рис. 3.15.

4 Цифровой спектральный анализ

4.1. Задачи и особенности цифрового спектрального анализа

Цифровой спектральный анализ (ЦСА) — это совокупность разнообразных методов обработки дискретных (цифровых) сигналов, позволяющих оценить частотный состав (спектр) анализируемого сигнала. Термин «цифровой спектральный анализ» в этом случае указывает только на то, что для спектрального анализа дискретных и цифровых сигналов применяются цифровые средства обработки.

В задачах ЦСА считается априорно известным, что анализируемый дискретный сигнал представляет собой сумму гармонических составляющих и, возможно, белого шума:

$$x(nT) = \sum_{j=0}^J A_j \cos(\omega_j nT + \varphi_j) + a_{\text{ш}} u(nT), \quad (4.1)$$

где A_j, ω_j, φ_j — параметры j -й гармонической составляющей; $a_{\text{ш}}$ — коэффициент включения белого шума (при наличии шума $a_{\text{ш}} = 1$, при его отсутствии $a_{\text{ш}} = 0$).

Естественно, что случайность параметров (параметра) гармонических составляющих в выражении (4.1) переводит анализируемый сигнал в класс случайных процессов даже при отсутствии шумового компонента.

Определение цифрового спектрального анализа, приведенное выше, дает достаточно широкое толкование предназначения этой процедуры. На практике ЦСА используется для решения большого числа разнообразных задач, имеющих только им присущие особенности и отличающихся своими целями. Очевидно, что различные задачи требуют обеспечения тех или иных условий реализации ЦСА. Практика цифровой обработки сигналов позволяет выделить три основные задачи цифрового спектрального анализа:

- спектральное оценивание;
- спектральное обнаружение;
- оценивание параметров спектральной составляющей.

Задача *спектрального оценивания* состоит в получении оценки спектра исследуемого сигнала в определенном диапазоне частот, что предполагает оценивание некоторой *функции частоты*. Спектр сигнала $x(nT)$, представленного выражением (4.1), на каждой j -й частоте будет определяться как гармоникой данной частоты, так и спектральной плотностью мощности белого шума $u(nT)$.

Задача *спектрального обнаружения* нацелена на принятие решения о наличии или отсутствии в смеси сигнала и шума полезной спектральной составляющей конкретной частоты. По отношению к выражению (4.1) спектральное обнаружение предназначено определить присутствие гармоники $A_j \cos(\omega_j nT + \varphi_j)$ на априорно известной частоте ω_j .

В некоторых случаях основной целью анализа является исследование локального поведения спектральной оценки в заданной узкой полосе частот или на конкретной частоте. Такая ситуация характерна для задачи *оценивания параметров спектральной составляющей*, в которой предметом оценивания могут быть амплитуда A_j , частота ω_j или начальная фаза φ_j спектральной составляющей.

Существуют ситуации, когда абсолютно точное решение любой из перечисленных задач не вызывает затруднений. Действительно, при анализе детерминированного финитного дискретного сигнала применение процедуры ДПФ, позволяющей определить амплитудный и фазовый спектры сигнала, обеспечивает достоверное решение любой задачи по цифровому спектральному анализу. На практике, как правило, условия выполнения ЦСА не соответствуют приведенному примеру. В этом случае успешное выполнение задач ЦСА может быть обеспечено только с учетом особенностей, характерных для цифрового спектрального анализа. Ниже представлены основные из них, значительно влияющие на выбор методов ЦСА и его эффективность.

Во-первых, существует необходимость максимального применения имеющихся сведений об анализируемом сигнале (например, учет известной величины J , известных частот ω_j или амплитуд A_j и т. п.). Правильное *использование априорной информации* об анализируемом сигнале создает наиболее благоприятные условия для повышения достоверности получаемых спектральных оценок при ограниченных ресурсных затратах.

Во-вторых, существенное влияние на результат ЦСА оказывает *частотное разрешение* Δf , обеспечиваемое применяемым мето-

дом анализа. Величина частотного разрешения характеризует минимальный частотный интервал между соседними спектральными составляющими, потенциально определяемыми системой ЦСА. Очевидно, что высокое (хорошее) частотное разрешение соответствует малым величинам этого интервала, а низкое (плохое) — значительным величинам. Следует отметить, что во многих видах практического применения частотное разрешение, обеспечиваемое тем или иным методом ЦСА, оценивается на основе визуальных впечатлений.

Различные задачи ЦСА предъявляют разные требования к частотному разрешению. Естественно, что наиболее требовательной к этому параметру является задача по спектральному обнаружению. И наоборот, задача по спектральному оцениванию, заключающаяся, как правило, в получении достаточно сглаженного спектра сигнала, менее требовательна к частотному разрешению. Из этой особенности следует важный для практики вывод о том, что не всегда хороший алгоритм спектрального оценивания может успешно применяться в задаче спектрального обнаружения.

В-третьих, в большинстве случаев анализируемый сигнал $x(nT)$ является случайным процессом. В такой ситуации простое вычисление ДПФ не представляет большого интереса, так как его результат содержит информацию лишь о спектре конкретной реализации случайного процесса. Спектральная оценка всего случайного сигнала характеризуется спектральной плотностью мощности (СПМ), расчет которой требует выполнения какого-либо усреднения результатов: статистического в общем случае или временного — в случае эргодичности анализируемого сигнала $x(nT)$.

В-четвертых, распространенной особенностью практического ЦСА является необходимость получения спектральных оценок бесконечного или достаточно длительного дискретного сигнала *по конечному числу отсчетов*, доступных для анализа. При большой длительности обрабатываемого сигнала оценка его спектрального состава производится посредством последовательного во времени расчета кратковременного спектра наблюдаемой финитной реализации, содержащей N отсчетов. Очевидно, что в этих условиях одним из основных факторов, определяющих точность ЦСА, является длительность интервала наблюдения t_n . Необходимо учитывать, что спектральная оценка, получаемая по конечному сегменту сигнала, характеризует лишь некоторое предположение относительно той истинной спектральной функции, которая была бы получена при доступности исследователю сигнала бесконечной длины. Из этого следует, что практика применения методов ЦСА бесконечных сигналов с исполь-

зованием их конечных сегментов в большинстве случаев базируется на эмпирическом опыте, а не на солидной теоретической основе.

С учетом отмеченных особенностей следует считать наиболее типичной задачу по цифровому спектральному анализу бесконечного или значительно длительного дискретного случайного сигнала $x(nT)$ в условиях доступности для обработки конечного числа отсчетов этого сигнала (конечного сегмента сигнала). При этом основными требованиями, предъявляемыми к системам ЦСА, выступают необходимое частотное разрешение и достаточная достоверность получаемого результата.

С целью демонстрации возможностей различных методов ЦСА рассмотрен дискретный тестовый сигнал

$$\text{test}(n) = \cos(0,275\pi n) + \cos(0,2\pi n) + 1,3u(n),$$

показанный на рис. 4.1 и определяемый линейной комбинацией двух гармонических составляющих и масштабированного шумового компонента $u(n)$, представляющего собой дискретный шум с нормальным распределением, нулевым математическим ожиданием и единичной дисперсией.

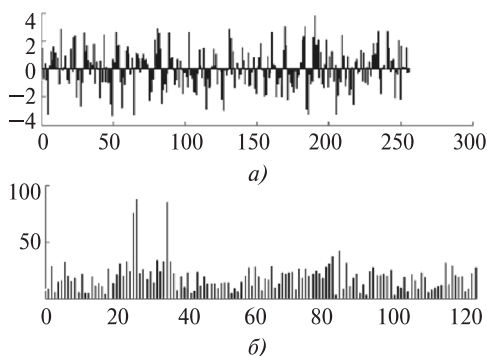


Рис. 4.1. Временное и спектральное представления финитной реализации тестового сигнала

Временное представление реализации тестового сигнала показано на рис. 4.1,а в виде конечного N -точечного сегмента ($N = 256$), изображенного на шкале нормированного времени n .

Спектральное представление реализации тестового сигнала (рис. 4.1,б) показано в виде половины периода модуля ДПФ сигнала, рассчитанного алгоритмом БПФ и показанного на шкале нормированной частоты $k = 0, 1, 2, \dots, N/2 - 1$.

Очевидно, что разность между значениями частот гармонических составляющих ($f_0 = 0,275$ и $f_0 = 0,2$) определяет *частотное разрешение* Δf , необходимое для корректного спектрального анализа сигнала $\text{test}(n)$. Наличие в спектре ложного гармонического компонента объясняется явлением растекания спектра, которое будет рассмотрено далее.

4.2. Классификация методов цифрового спектрального анализа

В современной практике цифровой обработки сигналов используется большое разнообразие методов ЦСА, что обуславливает необходимость системы их классификации. Характерно, что различные источники предлагают разные варианты классификации. В варианте классификации, показанном на рис. 4.2, за основу взята система, предложенная в [20]; авторские изменения и дополнения учитывают ряд традиционных для ЦСА методов и определений.



Рис. 4.2. Классификация методов цифрового спектрального анализа

Из рисунка видно, что все разнообразие способов ЦСА может быть представлено двумя классами методов: непараметрического и параметрического ЦСА. Классификационным признаком такого разделения выступает способ получения информации о спектральных характеристиках анализируемого сигнала.

В методах *непараметрического ЦСА* для получения спектральной оценки используется только информация, заключенная в анализируемых отсчетах сигнала, без каких-либо дополнительных предположений.

Совершенно другой подход предполагают методы *параметрического ЦСА*, в которых анализируемый сигнал отождествляется с некоторой настраиваемой математической моделью случайного процесса. ЦСА в данном случае сводится к решению оптимизационной задачи, заключающейся в определении таких *параметров* модели, при которых она максимально близка к анализируемому сигналу (достаточно часто такая процедура обозначается термином «настройка модели»). Очевидно, что спектральные оценки в методах параметрического ЦСА выражаются через параметры используемых моделей. Следует отметить, что в некоторых изданиях для обозначения параметрического ЦСА используется термин «модельный спектральный анализ» (MBSA — Model-Based Spectrum Analysis).

В теории и практике ЦОС используется более широкий набор методов цифрового спектрального анализа, чем представленный на рис. 4.2. Например, методы гармонического разложения Писаренко, анализа собственных векторов (EV — Eigenvectors) и классификации множественных сигналов (MUSIC — Multiple Signal Classification), которые в задачах спектрального оценивания, как правило, не используются, а предназначены для оценки параметров гармонических и узкополосных спектральных компонентов сигнала $x(nT)$.

4.3. Непараметрический цифровой спектральный анализ

Непараметрический ЦСА осуществляется двумя классами методов (см. рис. 4.2): на основе дискретного преобразования Фурье, которые часто в научной литературе называются классическими, и методами цифровой фильтрации. Такая классификация подразумевает, что в первом случае доступные для анализа отсчеты дискретного сигнала $x(nT)$ подвергаются процедуре вычисления ДПФ (следовательно, решается задача по спектральному оцениванию), а во втором случае задачи ЦСА решаются посредством обработки анализируемого сигнала цифровым фильтром (цифровыми фильтрами) со специальными характеристиками.

Необходимо отметить, что указанные классы методов непараметрического ЦСА достаточно близки друг к другу по выполняемым вычислительным операциям. Известно, что ДПФ может быть представлено как обработка дискретного сигнала набором цифровых фильтров, а дискретная фильтрация — с помощью процедуры ДПФ. Однако аспекты реализации ЦСА позволяют разделять используемые методы непараметрического анализа на два указанных класса.

4.3.1. Цифровой спектральный анализ методами цифровой фильтрации

Возможность использования цифровых фильтров для решения задачи по *спектральному обнаружению* следует из известных положений теории радиотехнической фильтрации, обосновывающих способы формирования фильтров с требуемыми характеристиками. Так, например, любой узкополосный цифровой фильтр с полосой пропускания ($f_{пп1} - f_{пп2}$) может являться инструментом для решения задачи по спектральному обнаружению гармонических составляющих с частотами от $f_{пп1}$ до $f_{пп2}$. Аналогичный вариант реализации устройства спектрального обнаружения может быть основан на использовании фильтра, согласованного с сигналом, в отношении которого решается задача ЦСА.

Для решения задачи по *спектральному оцениванию* одновременно используются несколько узкополосных ЦФ (УПЦФ) со смежными полосами пропускания. Спектральное оценивание на основе цифровой фильтрации (рис. 4.3) сводится к тому, что исследуемый сигнал $x(nT)$ в анализируемой области частот делится на R узких спектральных полос (подсигналов).

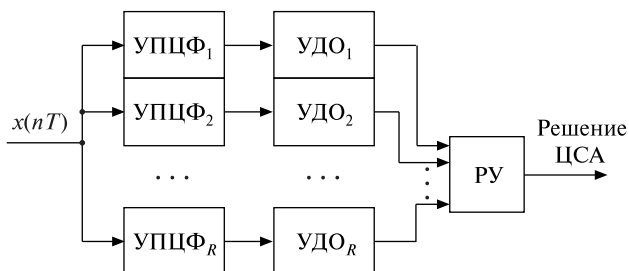


Рис. 4.3. Структурная схема системы ЦСА методом цифровой фильтрации

Выходной сигнал каждого узкополосного цифрового фильтра подается на устройство дополнительной обработки (УДО), в котором вычисляются средние значения амплитуд соответствующих подсигналов. На основании полученных значений решающее устройство (РУ) формирует решение о спектральной оценке анализируемого

сигнала. Точность ее зависит от избирательных свойств УПЦФ, в том числе от ширины полосы пропускания. Очевидно, что специальный алгоритм функционирования решающего устройства легко позволяет сформировать на основе системы, показанной на рис. 4.3, систему спектрального обнаружения одной или нескольких гармонических составляющих сигнала $x(nT)$.

4.3.2. Цифровое спектральное оценивание на основе дискретного преобразования Фурье

В большинстве случаев дискретные случайные сигналы, обрабатываемые в системах ЦОС, являются нестационарными. Расчет характеристик таких сигналов достаточно сложен, а в некоторых случаях и невозможен. В задачах ЦСА на основе дискретного преобразования Фурье полагается, что анализируемые сигналы стационарны и эргодичны. Такое допущение вполне корректно для практического применения (например, короткие сегменты сигнала, доступные для анализа, в большинстве случаев являются квазистационарными).

В теории и практике ЦОС используются два основных класса методов цифрового спектрального оценивания на основе ДПФ (см. рис. 4.2): с использованием периодограмм и с применением коррелограмм. Принципиальное различие между ними состоит в подходе к определению спектральной плотности мощности эргодического дискретного сигнала.

В методах на основе *периодограмм* (в некоторых источниках их называют *прямыми*) СПМ оценивается непосредственно по анализируемой дискретной последовательности. Для того чтобы по отсчетам дискретного сигнала получить спектральную оценку в единицах измерения мощности (СПМ), необходимо использовать несколько иную форму записи преобразований Фурье, чем те, которые используются в большинстве учебных пособий. С учетом введения в преобразование Фурье зависимости от масштабирующего коэффициента периода дискретизации T , непрерывный периодический спектр анализируемой бесконечной реализации случайного дискретного сигнала $x(nT)$ определяется прямым преобразованием Фурье в виде

$$X(e^{j\omega T}) [B, c] = T \sum_{n=-\infty}^{\infty} x(nT) e^{-j\omega nT} \quad \text{при} \quad -\frac{\pi}{T} \leq \omega \leq \frac{\pi}{T}. \quad (4.2)$$

Квадрат модуля спектра преобразования (4.2) характеризует распределение энергии сигнала $x(nT)$ по частоте:

$$|X(e^{j\omega T})|^2 [B^2, c^2] [B_T, c] = \left| T \sum_{n=-\infty}^{\infty} x(nT) e^{-j\omega nT} \right|^2$$

при $-\pi/T \leq \omega \leq \pi/T$.

Так как мощность сигнала равна его энергии, отнесенной к единице времени, то распределение мощности анализируемой реализации сигнала $x(nT)$ по частоте $G_p(\omega)$ определяется следующим выражением:

$$G_p(\omega)[\text{Вт}] = \frac{1}{NT} \left| T \sum_{n=-\infty}^{\infty} x(nT) e^{-j\omega nT} \right|^2 = \\ = \frac{T}{N} \left| \sum_{n=-\infty}^{\infty} x(nT) e^{-j\omega nT} \right|^2 \quad \text{при } -\frac{\pi}{T} \leq \omega \leq \frac{\pi}{T}.$$

Естественно, что СПМ $G(\omega)$ всего ансамбля реализаций выражается через математическое ожидание функции $G_p(\omega)$:

$$G(\omega) = M\{G_p(\omega)\}.$$

В случае предельной длительности анализируемого сигнала при $N \rightarrow \infty$ искомая спектральная мощность определяется следующим образом:

$$G(\omega) = \lim_{N \rightarrow \infty} M \left\{ \frac{T}{2N+1} \left| \sum_{n=-N}^N x(nT) e^{-j\omega nT} \right|^2 \right\} \quad \text{при } -\frac{\pi}{T} \leq \omega \leq \frac{\pi}{T}.$$

Допущение об эргодичности анализируемого сигнала $x(nT)$ позволяет отказаться от процедуры статистического усреднения, так как все реализации эргодического процесса имеют одинаковые спектры мощности. Отказ от вычисления математического ожидания и учет конечности наблюдаемого сигнала ($n = 0, 1, 2, \dots, N-1$) позволяют получить конечное выражение для расчета дискретной СПМ, вычисляемой по конечному числу отсчетов сигнала на основе дискретного преобразования Фурье:

$$G(k\Omega) = \frac{T}{N} \left| \sum_{n=0}^{N-1} x(nT) e^{-jk\Omega nT} \right|^2 \quad \text{при } \Omega = \frac{2\pi}{NT}. \quad (4.3)$$

Выражение (4.3), показывающее способ оценки СПМ случайного эргодического сигнала $x(nT)$ по N отсчетам его одной реализации, получило название *периодограмма*, а метод спектрального оценивания на основе формулы (4.3) — *метод периодограммы*.

Различные варианты использования изложенного выше подхода к определению СПМ эргодического дискретного сигнала обуславливают наличие нескольких методов цифрового спектрального оценивания на основе периодограмм (см. рис. 4.2).

Если при оценке СПМ используется оконная функция $p(n)$ при $n = 0, 1, 2, \dots, N - 1$, выражение (4.3) преобразуется к виду

$$G(k\Omega) = \frac{T}{\sum_{n=0}^{N-1} |p(n)|^2} \left| \sum_{n=0}^{N-1} x(nT)p(n)e^{-jk\Omega nT} \right|^2 \quad (4.4)$$

и называется *модифицированной периодограммой*. Соответствующий метод оценки СПМ носит название *метод модифицированной периодограммы*.

Особенностью периодограмм (4.3) и (4.4) является их значительная изрезанность при вычислении по длительному сегменту реализации случайного сигнала, что не всегда желательно. Для уменьшения изрезанности оценки СПМ в практике ЦСА применяются различные способы сглаживания периодограмм.

В методе *Даньелла* сглаживание быстрых флуктуаций СПМ достигается усреднением выражения (4.3) по соседним спектральным частотам. Периодограмма Даньелла $G_D(k\Omega)$ рассчитывается посредством усреднения спектральных составляющих периодограммы (4.3) в P точках с каждой стороны от частоты $k\Omega$:

$$G_D(k\Omega) = \frac{1}{2P+1} \sum_{k-P}^{k+P} G(k\Omega).$$

В методе *Барлетта* анализируемый сигнал $x(nT)$ делится на неперекрывающиеся сегменты, для каждого из которых в соответствии с выражением (4.3) вычисляется частная периодограмма, после чего искомая оценка СПМ определяется усреднением по множеству частных периодограмм.

Метод *Уэлча* (усреднения модифицированных периодограмм) отличается от метода *Барлетта* двумя усовершенствованиями: разделением анализируемого сигнала $x(nT)$ на перекрывающиеся сегменты и использованием оконных функций для расчета частных периодограмм сегментов. Применение оконных функций в методе Уэлча направлено на ослабление эффекта боковых лепестков (эффекта растекания спектра), а целью перекрытия сегментов является увеличение числа усредняемых частных периодограмм при сохранении значения N , что обеспечивает уменьшение дисперсии оценки СПМ. Детальное представление алгоритма цифрового спектрального оценивания методом Уэлча изложено в разд. 4.3.4.

В методах на основе *коррелограмм* (в некоторых источниках такие методы называются *косвенными*) применяется предварительное определение автокорреляционной функции (АКФ) анализируемого

случайного дискретного сигнала $x(nT)$:

$$B(m) = M\{x(n+m)x(n)\},$$

которая в случае эргодичности $x(nT)$ может быть рассчитана не статистическим, а временным усреднением по единственной реализации сигнала. В случае бесконечной длительности анализируемого сигнала ($N \rightarrow \infty$) соответствующее выражение имеет вид

$$B(m) = M\{x(n+m)x(n)\} = \lim_{N \rightarrow \infty} \left\{ \frac{1}{2N+1} \sum_{n=-N}^N x(n+m)x(n) \right\}.$$

СПМ случайного дискретного сигнала $x(nT)$ в соответствии с теоремой Винера–Хинчина определяется прямым преобразованием Фурье автокорреляционной функции:

$$G(\omega) = T \sum_{m=-\infty}^{\infty} B(m)e^{-j\omega mT} \quad \text{при} \quad -\frac{\pi}{T} \leq \omega \leq \frac{\pi}{T}. \quad (4.5)$$

Методы оценивания СПМ на основе коррелограмм подразумевает подстановку в выражение (4.5) конечной последовательности значений АКФ. В частном случае ($m = 0, 1, 2, \dots, N-1$) дискретная СПМ определяется дискретным преобразованием Фурье по N точкам АКФ единственной реализации:

$$G(k\Omega) = T \sum_{m=0}^{N-1} B(m)e^{-jk\Omega mT} \quad \text{при} \quad \Omega = \frac{2\pi}{NT}. \quad (4.6)$$

Выражение (4.6) получило название *коррелограмма*. Различные особенности использования коррелограммы для расчета СПМ анализируемого случайного сигнала обуславливают наличие разных методов цифрового спектрального оценивания на основе коррелограмм. Например, использование оконной функции $p(m)$ формирует метод *модифицированной коррелограммы*:

$$G(k\Omega) = T \sum_{m=0}^{N-1} B(m)p(m)e^{-jk\Omega mT} \quad \text{при} \quad \Omega = \frac{2\pi}{NT}.$$

Классические методы *комбинированного* спектрального оценивания используют различные варианты объединения достоинств методов на основе периодограмм и коррелограмм.

4.3.3. Оконные функции

Классические методы ЦСА реализуют процедуру спектрального оценивания по результатам последовательной обработки конечных кадров исходного сигнала. Применение ДПФ для задач ЦСА

требует перехода от исходного сигнала $x(n)$ или его автокорреляционной функции, имеющей (-его) бесконечную длительность, к финитным сегментам (кадрам) обработки, содержащим N отсчетов. При этом обрабатываемый конечный сегмент сигнала или АКФ удобно рассматривать как некоторую часть соответствующей бесконечной последовательности, видимую через применяемое окно. Такое преобразование математически определяется операцией взвешивания последовательности $x(n)$ при $n = 0, 1, 2, \dots$ или $B(m)$ при $m = 0, 1, 2, \dots$ оконными функциями (функциями взвешивания) $p(n)$ при $n = 0, 1, 2, \dots, N - 1$ (окно данных) или $p(m)$ при $m = 0, 1, 2, \dots, N - 1$ (корреляционное окно) соответственно. Необходимо отметить, что в качестве корреляционных окон используются только окна с нечетными N , поскольку точка симметрии таких окон приходится на средний отсчет $p(m)$, равный единице.

В методах на основе периодограмм выражение для расчета ДПФ взвешенного сигнала имеет вид

$$X_p(k) = \sum_{n=0}^{N-1} x(n)p(n)e^{-j(2\pi/N)kn} = \sum_{n=0}^{N-1} x_p(n)e^{-j(2\pi/N)kn}, \quad (4.7)$$

где $x_p(n) = x(n)p(n)$ — взвешенный сигнал; $k = 0, 1, 2, \dots, N - 1$.

В простейшем случае сигнал $x_p(n)$ формируется путем выделения конечной N -точечной выборки (обрабатываемого сегмента) из сигнала $x(n)$

$$x_p(n) = \begin{cases} x(n) & \text{при } n = 0, 1, 2, \dots, N - 1; \\ 0 & \text{при } n < 0, n > N - 1, \end{cases}$$

при этом используется прямоугольная оконная функция (ОФ):

$$p(n) = \begin{cases} 1 & \text{при } 0 \leq n \leq N - 1; \\ 0 & \text{при } n < 0, n > N - 1. \end{cases}$$

Очевидно, что такой подход не обеспечивает точного результата спектрального оценивания исходного бесконечного сигнала. Действительно, оценка спектра, получаемая на основании выражения (4.7), предполагает, что обрабатываемый сегмент периодически повторяется во времени, что на самом деле не имеет места. Если значения начального и конечного отсчетов обрабатываемого сегмента при этом отличаются значительно, при периодическом повторении обрабатываемых сегментов на их стыках возникают скачки, обуславливающие расширение спектра сигнала. Такое явление, вызванное переходом от сигнала бесконечной длительности к финитному сегменту обработки, в литературе часто называется *растеканием спектра*.

Иллюстрация *растекания спектра* показана на рис. 4.4 и 4.5 на примере цифрового спектрального оценивания гармонического дис-

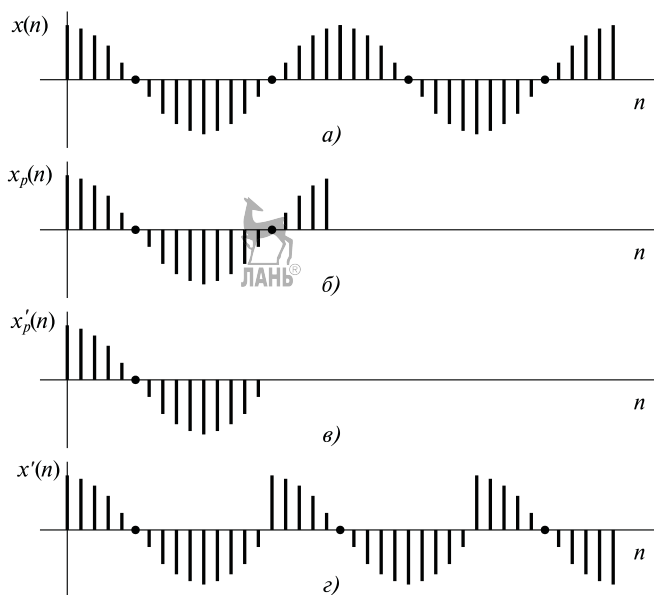


Рис. 4.4. Иллюстрация процедуры формирования обрабатываемого сегмента на основе прямоугольной оконной функции

кретного сигнала. Анализируемый сигнал $x(nT) = A \cos(\omega nT)$ представлен на рис. 4.4,а.

Если обрабатываемый сегмент конечной длины $x_p(nT)$ содержит целое число периодов исходного сигнала $x(nT)$, то результат спектрального оценивания $x_p(nT)$ точно характеризует спектральный состав сигнала $x(nT)$. Периодическое продолжение $x_p(nT)$ приводит к формированию исходного сигнала $x(nT)$, что обуславливает корректность выполнения процедуры ДПФ. Такой случай показан на рис. 4.4,б (обрабатываемый сегмент отображает один период анализируемого сигнала).

На рис. 4.4,в представлен случай, когда обрабатываемый сегмент $x'_p(nT)$ содержит нецелое число периодов анализируемого гармонического сигнала $x(nT)$. Очевидно, что в этом случае по условиям выполнения ДПФ результат спектрального оценивания будет отображать спектр сигнала $x'(nT)$, формируемого периодическим повторением сегмента $x'_p(nT)$ и показанного на рис. 4.4,г. Сигнал $x'(nT)$ не является гармоническим, следовательно, его спектр шире спектра анализируемого сигнала $x(nT)$. Происшедшее *растекание спектра* обуславливает неточность спектральной оценки сигнала $x(nT)$ по сегменту $x'_p(nT)$.

На рис. 4.5,а и в представлены обрабатываемые сегменты $x_p(n)$

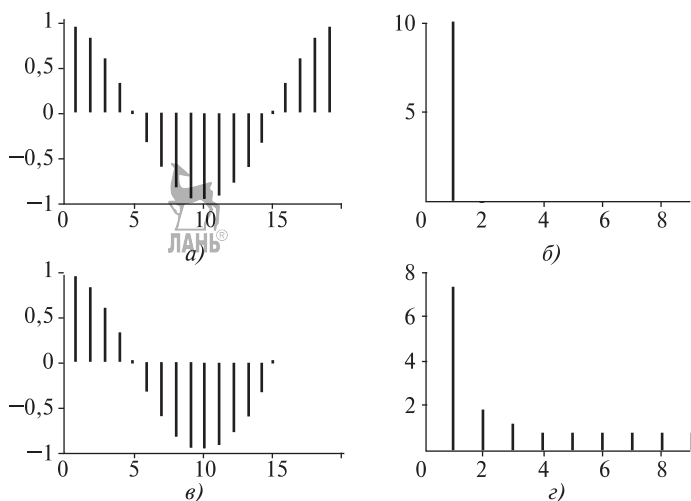


Рис. 4.5. Иллюстрация растекания спектра при цифровом спектральном оценивании дискретного гармонического сигнала на основе ДПФ

и $x'_p(n)$ соответственно для единичной амплитуды. Результаты их спектрального оценивания, показанные на рис. 4.5, б и г, отображают половины периодов соответствующих амплитудных спектров $|X_p(k)|$ и $|X'_p(k)|$ и наглядно демонстрируют явление растекания спектра.

Естественно, что аналогичные процессы, обуславливающие растекание спектра, обязательно будут наблюдаться и при оценке сложных, в общем случае непериодических сигналов.

Взвешивание исходного сигнала во временной области приводит к изменению и его спектральных характеристик. Из свойств преобразования Фурье известно, что спектр произведения двух дискретных сигналов $x_p(n) = x(n)p(n)$ представляет собой круговую свертку их спектров. Следовательно, выбор оконной функции оказывает значительное влияние на результат цифрового спектрального оценивания, который зависит от спектральных характеристик используемой оконной функции. Следует отметить, что фазовые спектры всех ОФ, представленных в этом разделе, являются линейными функциями частоты и по этой причине не рассматриваются. Амплитудные спектры оконных функций определяются модулями преобразований Фурье от временных зависимостей $p(n)$.

Временное и спектральное представления прямоугольной ОФ показаны на рис. 4.6. Во временной области (рис. 4.6, а) прямоугольная ОФ изображена в виде $p(n-1)$, соответствующем сдвигу вправо на один отсчет. В спектральной области (рис. 4.6, б) изображена половина периода четного амплитудного спектра $|P(e^{j2\pi f_0})|$

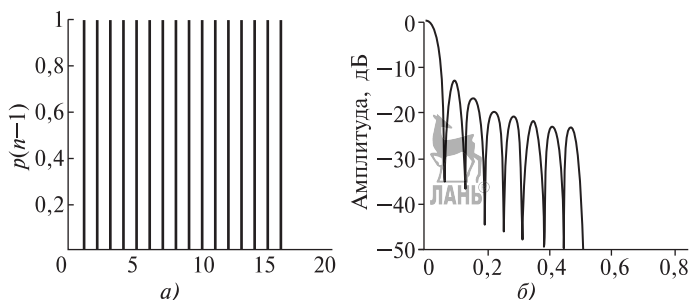


Рис. 4.6. Временное (а) и спектральное (б) представления прямоугольного окна

прямоугольной оконной функции $p(n)$, $n = 0, 1, 2, \dots, 15$, по нормированной частоте f_0 , при этом в качестве единиц измерения амплитуды использованы децибелы. Для обеспечения на нулевой частоте единичного значения амплитуды ($P = 0$ дБ) использовано предварительное нормирование окна: перед расчетом спектра отсчеты $p(n)$ уменьшены в число раз, равное их сумме (в 16 раз).

Из рис. 4.6 понятно, что минимальная ширина спектральных пиков взвешенного сегмента $x_p(nT)$ ограничена шириной главного лепестка амплитудного спектра прямоугольного окна и не зависит от значений $x_p(nT)$. Боковые лепестки $|P(e^{j2\pi f_0})|$ изменяют амплитуды соседних спектральных пиков анализируемого сигнала. Уровень первого бокового лепестка амплитудного спектра прямоугольной ОФ значителен и составляет $-13,3$ дБ.

Различные задачи ЦСА формируют различные требования к оконным функциям. В теории и практике цифрового спектрального оценивания на основе ДПФ используется широкий перечень непрямоугольных оконных функций, некоторые из которых представлены в табл. 4.1, при этом диапазон изменения нормированного времени определяется областью $n = 0, 1, 2, \dots, N-1$.

Основными особенностями непрямоугольных ОФ во временной области являются наличие максимума в середине окна (при $n = (N-1)/2$ для нечетных N) и плавное спадание к краям (при $n = 0$ и $n = N-1$), что обеспечивает ослабление нежелательных эффектов, связанных с возникновением скачков сигнала при периодическом повторении обрабатываемого сегмента $x_p(nT)$. Естественно, что такой подход позволяет ослабить явление растекания спектра.

В частотной области используемые непрямоугольные ОФ характеризуются меньшими уровнями боковых лепестков по сравнению с прямоугольным окном, что обеспечивает повышение точности спектрального оценивания. Однако это достигается ценой расширения главных лепестков оконных функций, что приводит к ухудше-

Таблица 4.1

Некоторые ОФ и их характеристики для $n = 0, 1, 2, \dots, N - 1$

Оконная функция	Аналитическое выражение $p(n)$ оконной функции	$\Delta F_{\text{ш}}$, бин	M_6 , дБ	V_6 , дБ/октава
Прямоугольное окно	1	1,00	-13,3	-6
Треугольное окно	$1 - \frac{ 2n - N + 1 }{N - 1}$	1,33	-26,5	-12
Окно Хэмминга	$0,54 + 0,46 \cos \left[\frac{\pi(2n - N + 1)}{N - 1} \right]$	1,36	-43	-6
Окно Гаусса	$\exp \left[-12,5 \left(\frac{n - \frac{N-1}{2}}{N - 1} \right)^2 \right]$	1,39	-42	-6
Окно Ханна	$0,5 + 0,5 \cos \left[\frac{\pi(2n - N + 1)}{N - 1} \right]$	1,5	-31,5	-18
Окно Натолла	$\sum_{r=0}^3 a_r \cos \left[\frac{\pi r(2n - N + 1)}{N - 1} \right],$ где $a_0 = 0,3635819,$ $a_1 = 0,4491775, a_2 = 0,1365995,$ $a_3 = 0,0106411$	1,8	-98	-6

нию частотного разрешения процедуры ЦСА. Следовательно, выбор используемой ОФ должен определяться на основании компромисса между шириной главного лепестка и уровнем подавления боковых.

Для оценки качества оконных функций используется множество характеристик. Ниже представлены характеристики ОФ, наиболее существенные для решения задач спектрального анализа и используемые в табл. 4.1.

Частотное разрешение, обеспечиваемое классическим методом спектрального оценивания при использовании оконной функции, определяется в основном шириной полосы частот главного лепестка амплитудного спектра ОФ. В теории и практике ЦСА для качественной оценки ширины полосы главного лепестка принято использовать эквивалентную шумовую полосу (ЭШП) $\Delta F_{\text{ш}N}$ ОФ, измеряемую в бинах ДПФ следующим образом:

$$\Delta F_{\text{ш}N} = \frac{N \sum_{n=0}^{N-1} p^2(n)}{\left[\sum_{n=0}^{N-1} p(n) \right]^2}.$$

Как правило, $N \gg 1$ и вместо $\Delta F_{\text{ш}N}$ используется предельная ЭШП $\Delta F_{\text{ш}}$, определяемая при условии стремления N к бесконечности: $\Delta F_{\text{ш}} = \lim_{N \rightarrow \infty} \Delta F_{\text{ш}N}$.

Эквивалентная шумовая полоса ОФ $\Delta F_{\text{ш}}$ соответствует удвоенной ширине полосы пропускания идеального ФНЧ, у которого максимальное значение АЧХ равно максимальному значению амплитудного спектра оконной функции, а мощность шума, пропускаемого фильтром, равна мощности шума после обработки его оконной функцией.

Для качественной оценки боковых лепестков амплитудного спектра оконной функции используются две характеристики. *Максимальный уровень M_6 боковых лепестков*, измеряемый в децибелах по отношению к уровню главного лепестка, и *асимптотическая скорость V_6 спада боковых лепестков*, измеряемая в децибелах на октаву изменения ширины полосы частот, характеризуют степень «просачивания» «лишних» спектральных составляющих при решении задачи по спектральному оцениванию. Естественно, повышение точности такого оценивания обеспечивается при уменьшении M_6 и возрастании V_6 .

Анализ содержания табл. 4.1 свидетельствует о том, что самый узкий главный лепесток из ОФ, приведенных в таблице, имеет амплитудный спектр прямоугольного окна, но это же окно имеет и самый высокий уровень боковых лепестков. Противоположно предельные характеристики имеет окно Натолла: при наименьшем максимальном уровне боковых лепестков эта ОФ имеет самый широкий главный лепесток амплитудного спектра.

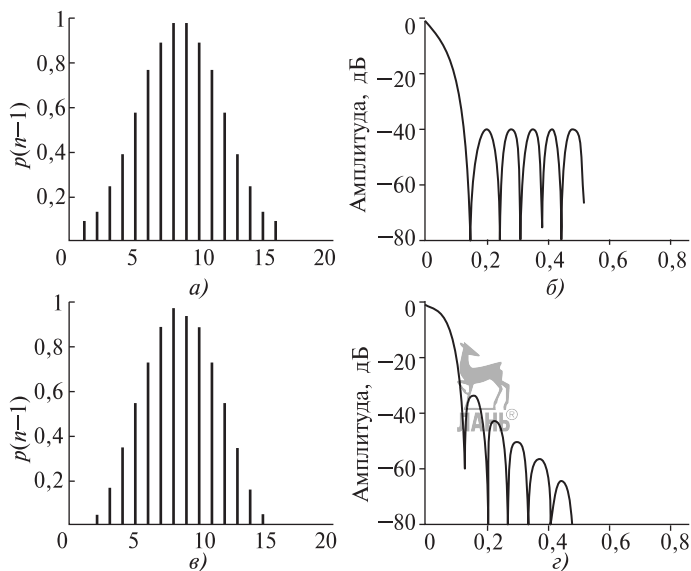


Рис. 4.7. Временное и спектральное представления окон Хэмминга и Ханна

Рис. 4.7 содержит временное и спектральное представления окна Хэмминга (рис. 4.7, *а* и *б*) и окна Ханна (рис. 4.7, *в* и *г*), причем графики построены согласно условий, указанным для прямоугольного окна (см. рис. 4.6). Графики амплитудных спектров наглядно подтверждают наличие большей асимптотической скорости спада боковых лепестков у окна Ханна при незначительном различии ширины главного лепестка.

Необходимо отметить, что существующий набор оконных функций не является полным. Достаточно часто для конкретного приложения ЦОС разрабатываются специализированные оконные функции, соответствующие требованиям решаемых задач по спектральному анализу.

4.3.4. Цифровое спектральное оценивание методом усреднения модифицированных периодограмм

Метод усреднения модифицированных периодограмм (метод Уэлча) является одним из популярных классических методов цифрового спектрального оценивания. Ниже представлены основные этапы оценки СПМ случайного сигнала данным методом.

Первый этап заключается в формировании исходных данных решаемой задачи ЦСА. На этом этапе определяются период дискретизации T исходного сигнала $x(nT)$ и требование к частотному разрешению Δf спектрального оценивания. Следствием является определение длительности достаточного интервала наблюдения t_n и числа отсчетов сигнала $x(nT)$, обрабатываемых на этом интервале:

$$N = \text{int} [t_n / T]. \quad (4.8)$$

Второй этап включает процедуру обработки сегмента сигнала на интервале наблюдения t_n и состоит из нескольких операций.

1. Дополнение числа отсчетов N , определенного на основании выражения (4.8), до целой степени числа 2 для обеспечения возможности применения стандартных алгоритмов БПФ. На практике это означает дополнение последовательности отсчетов сигнала $x(nT)$ на интервале наблюдения необходимым числом нулей для формирования анализируемого сегмента $x_\nu(nT)$ на ν -м интервале наблюдения $t_{n\nu}$.

2. Выбор оконной функции $p(n)$ в соответствии с требованиями решаемой задачи по спектральному оцениванию.

3. Вычисление ДПФ взвешенного сегмента сигнала алгоритмом БПФ на ν -м интервале наблюдения $t_{n\nu}$:

$$X_\nu(k) = \sum_{n=0}^{N-1} x_\nu(n)p(n)e^{-j(2\pi/N)kn} \quad \text{при } k = 0, 1, 2, \dots, N-1.$$

4. Вычисление модифицированной периодограммы $G_\nu(k)$ ν -го сегмента:

$$G_\nu(k) = \frac{T|X_\nu(k)|^2}{\sum_{n=0}^{N-1} |p(n)|^2}. \quad (4.9)$$

В случае невозможности принятия решения в реализуемой задаче цифрового оценивания по результатам расчета $G_\nu(k)$ следует для устранения неопределенности повторить операции 2, 3 и 4, увеличив N в 2^l раз, где l — целое число, при этом вместо формулы (4.9) используется модифицированная формула

$$X'_\nu(k) = \sum_{n=0}^{N2^l-1} x'_\nu(n)p(n)e^{-j(2\pi/N)2^l kn} \quad \text{при } k = 0, 1, \dots, N2^l - 1,$$

$$\text{где } x'_\nu(n) = \begin{cases} x(n) & \text{при } n = 0, 1, 2, \dots, N-1; \\ 0 & \text{при } n = N, N+1, N+2, \dots, N \cdot 2^l - 1. \end{cases}$$

Представленная модификация означает дополнение анализируемого сегмента $x_\nu(nT)$, содержащего N отсчетов, нулями до длительности в $N \cdot 2^l$ отсчетов. В результате количество спектральных составляющих анализируемого сигнала на периоде его спектра, определяемое ДПФ, возрастет с N до $N \cdot 2^l$, что означает повышение частотного разрешения ЦСА.

Третий этап заключается в статистическом усреднении результатов второго этапа, полученных для нескольких интервалов наблюдения, с целью уменьшения дисперсии полученной оценки СПМ. Ниже представлены основные операции третьего этапа.

1. Выбор коэффициента перекрытия соседних интервалов наблюдения D . Варианты достаточно популярных значений $D = 0,5$ и $D = 0,75$ представлены на рис. 4.8, а и б соответственно.

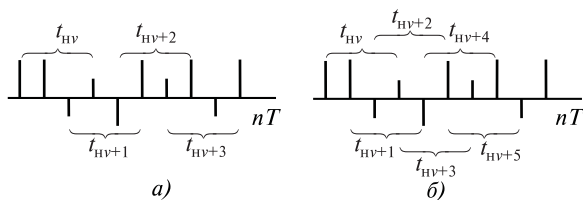


Рис. 4.8. Варианты перекрытия соседних интервалов наблюдения

2. Расчет доступного числа интервалов наблюдения:

$$V = \text{int} \left[\frac{L - DN}{N - DN} \right],$$

где L — общее число отсчетов сигнала, используемое в процедуре спектрального оценивания. Так, для рис. 4.8,а доступное число интервалов наблюдения составляет $V = \text{int} \left[\frac{10 - 0,5 \cdot 4}{4 - 0,5 \cdot 4} \right] = 4$.

3. Расчет усредненной оценки СПМ:

$$G(k) = \frac{1}{V} \sum_{v=1}^V G_v(k).$$

4.4. Параметрический цифровой спектральный анализ

Как было отмечено выше, параметрический цифровой спектральный анализ основан на использовании некоторой математической модели анализируемого случайного дискретного сигнала, описываемой конечным набором параметров. Процедура анализа в этом случае сводится к поиску таких значений параметров, при которых настраиваемая модель наиболее близка к реальному анализируемому сигналу $x(nT)$. Результатом параметрического ЦСА является выражение искомого результата (например, СПМ в случае спектрального оценивания) через параметры используемой модели.

Одним из преимуществ применения параметрических моделей случайных процессов является возможность получения на их основе более точных оценок СПМ $G(\omega)$ исследуемого сигнала, чем с помощью классических методов ЦСА. Другое значимое преимущество — возможность достижения более высокого спектрального разрешения. Классические методы ЦСА дают оценки СПМ по взвешенным последовательностям исходного сигнала или его автокорреляционной функции. Отсчеты сигнала или АКФ за пределами применяемого окна полагаются равными нулю, что, естественно, является достаточно грубым допущением и приводит к искажениям спектральных оценок. На практике часто имеется некоторая априорная информация относительно анализируемого сигнала, которую можно использовать для построения модели $y(nT)$, достаточно точно аппроксимирующей сигнал $x(nT)$. Полученные в этих условиях модели позволяют принимать более реалистические допущения о данных вне окна, чем допущение об их равенстве нулю. В результате отпадает необходимость в оконных функциях $p(n)$, а следовательно, устраняются и связанные с ними искажения. Степень улучшения разрешения и повышения достоверности спектральных оценок определяется соответствием выбранной модели анализируемому сигналу. Большие распространение параметрические методы ЦСА получили для решения задачи по цифровому спектральному оцениванию.

4.4.1. Принцип параметрического спектрального оценивания

Принцип параметрического спектрального оценивания проиллюстрирован на рис. 4.9.

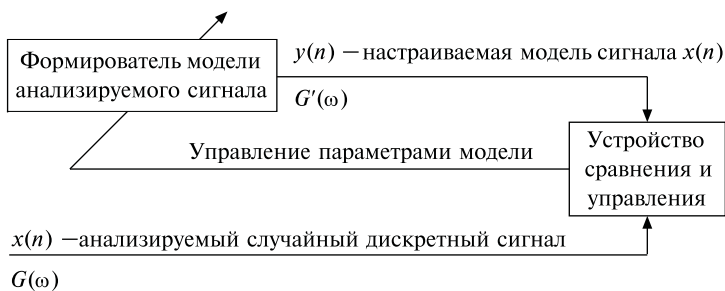


Рис. 4.9. Принцип параметрического спектрального оценивания

Анализируемый случайный дискретный сигнал $x(n)$, имеющий спектральную плотность мощности $G(\omega)$, поступает на вход устройства сравнения и управления (УСУ). На второй вход УСУ поступает случайный дискретный сигнал $y(n)$, моделирующий анализируемую последовательность и характеризующийся спектральной плотностью мощности $G'(\omega)$, выраженной через параметры формирователя (параметры модели). Функциональным предназначением УСУ являются сравнение сигналов $x(n)$ и $y(n)$ с использованием какого-либо критерия похожести и последующее управление выбором таких значений определенного заранее набора параметров, которые обеспечивают максимальное сходство модели и оригинала: $y(n) \approx x(n)$. Очевидно, что теоретически, без учета ресурсных ограничений, возможно достижение любой требуемой степени похожести модели и оригинала, что делает возможным представление спектральной плотности мощности анализируемого сигнала через СПМ модели: $G(\omega) \approx G'(\omega)$.

Основными этапами параметрического спектрального оценивания выступают:

- выбор параметрической модели временного ряда (типа настраиваемой модели), соответствующей имеющейся априорной информации об анализируемом сигнале $x(n)$;
- вычисление оценок параметров модели по результатам наблюдения имеющихся отсчетов анализируемого сигнала (параметрическая идентификация модели);
- расчет искомой оценки СПМ $G(\omega)$ анализируемого сигнала $x(n)$ на основе полученных значений параметров используемой модели.

Теория и практика ЦОС применяют широкий набор типов настраиваемых моделей, используемых для решения задач ЦСА. Все они являются классификационным признаком методов параметрического цифрового спектрального оценивания.

4.4.2. Параметрическое спектральное оценивание на основе моделей сигналов в виде реакции линейных цифровых фильтров

Методы параметрического спектрального оценивания основаны на том, что анализируемый сигнал $x(nT)$, представляющий собой сумму детерминированных и случайных дискретных сигналов, может быть достаточно хорошо аппроксимирован реакцией $y(nT)$ линейного цифрового фильтра (ЦФ), описываемого линейным разностным уравнением

$$y(nT) = - \sum_{m=1}^{M-1} a_m y(nT - mT) + \sum_{k=0}^{K-1} b_k u(nT - kT), \quad (4.10)$$

на дискретный сигнал $u(n)$. Существенно, что вещественные коэффициенты a_m и b_k могут принимать как положительные, так и отрицательные значения, поэтому в некоторых изданиях используются иные знаки при суммах правой части выражения (4.10).

В качестве воздействия $u(n)$, как правило, используются отсчеты белого шума с нулевым математическим ожиданием и дисперсией σ^2 . Дискретный белый шум, как известно, не коррелирован сам с собой при любых временных сдвигах, за исключением $m = 0$, а его АКФ имеет вид

$$B_u(m) = \sigma^2 \delta(m),$$

где $\delta(m)$ — цифровой единичный импульс. Спектральная плотность мощности дискретного белого шума определяется выражением

$$G_u(\omega) = \sigma^2 T.$$

Частотная характеристика ЦФ

$$H(e^{j\omega T}) = \frac{\sum_{k=0}^{K-1} b_k e^{-jk\omega T}}{1 + \sum_{m=1}^{M-1} a_m e^{-jm\omega T}},$$

модуль которой определяет АЧХ фильтра:

$$A(\omega T) = |H(e^{j\omega T})| = \sqrt{H_{\text{Re}}^2(e^{j\omega T}) + H_{\text{Im}}^2(e^{j\omega T})} =$$

$$= \sqrt{\frac{\left(\sum_{k=0}^{K-1} b_k \cos k\omega T\right)^2 + \left(\sum_{k=0}^{K-1} b_k \sin k\omega T\right)^2}{\left(1 + \sum_{m=1}^{M-1} a_m \cos m\omega T\right)^2 + \left(\sum_{m=1}^{M-1} a_m \sin m\omega T\right)^2}}.$$

Структура рассматриваемой процедуры цифрового спектрального оценивания показана на рис. 4.10.

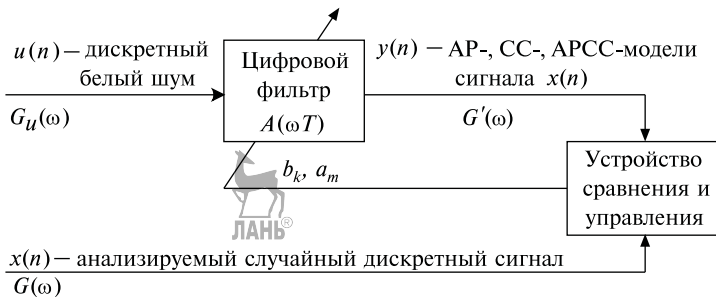


Рис. 4.10. Принцип параметрического спектрального оценивания на основе моделей сигналов в виде реакции линейных цифровых фильтров

Спектральная плотность мощности сигнала $y(n)$ определяется через параметры ЦФ и белого шума следующим образом:

$$G'(\omega) = G_u(\omega) A^2(\omega T) = \sigma^2 T \left[\frac{\left(\sum_{k=0}^{K-1} b_k \cos k\omega T\right)^2 + \left(\sum_{k=0}^{K-1} b_k \sin k\omega T\right)^2}{\left(1 + \sum_{m=1}^{M-1} a_m \cos m\omega T\right)^2 + \left(\sum_{m=1}^{M-1} a_m \sin m\omega T\right)^2} \right]. \quad (4.11)$$

Устройство сравнения и управления обеспечивает выбор значений коэффициентов b_k и a_m , при которых реакция ЦФ максимально близка к анализируемому дискретному сигналу. Степень точности аппроксимации исследуемого сигнала $x(n) \approx y(n)$ детерминирует точность получаемой оценки СПМ, выраженной на основе формулы (4.11) через параметры цифрового фильтра и дисперсию белого шумового воздействия: $G(\omega) \approx G'(\omega)$.

Различные ограничения на значения параметров a_m и b_k позволяют аппроксимировать анализируемый сигнал моделями авторегрессионного процесса (АР-моделью), процесса со скользящим средним (СС-моделью) и процесса авторегрессии — скользящего среднего (АРСС-моделью). Выбор конкретной модели из перечисленных требует некоторых предварительных сведений о возможной форме

СПМ исследуемого сигнала. Для моделирования спектров с острыми пиками, но без глубоких впадин, наиболее подходящей является авторегрессионная модель. Она наиболее применима и в случае наличия ограничений по производительности используемых цифровых процессоров обработки сигналов, так как является наименее затратной по вычислительным ресурсам. Для сигналов, спектры которых имеют глубокие впадины при отсутствии острых пиков, наиболее адекватна модель процесса со скользящим средним. АРСС-модель (как наиболее общая) может использоваться для аппроксимации любого исследуемого сигнала, однако она менее точна, чем АР-модель, при описании пиков СПМ и хуже, чем СС-модель, аппроксимирует провалы спектра.

Принципиально и то, что в соответствии с теоремой декомпозиции любой стационарный сигнал $x(n)$ с конечной дисперсией, описываемый одной из рассматриваемых моделей, может быть представлен любой из двух других моделей большого порядка.

4.4.3. Параметрическое спектральное оценивание на основе модели авторегрессионного процесса

Модель авторегрессионного процесса (AR — Autoregressive) получила наибольшее распространение среди параметрических моделей, используемых при решении задач по спектральному оцениванию дискретных случайных процессов. Это обусловлено простотой, удобством расчетов на ее основе и достаточной степенью ее адекватности большому классу реальных сигналов. Модель определяется частным случаем выражения (4.10) при ограничениях на значения параметров, масштабирующих отсчеты шумового воздействия:

$$b_k = \begin{cases} 1 & \text{при } k = 0; \\ 0 & \text{при } k \neq 0, \end{cases}$$

и имеет вид

$$y(nT) = - \sum_{m=1}^{M-1} a_m y(nT - mT) + u(nT). \quad (4.12)$$

В методах параметрического цифрового спектрального оценивания на основе АР-модели в качестве формирователя модели анализируемого сигнала выступает адаптивный РЦФ, порядок которого определяет точность аппроксимации сигнала $x(nT)$.

Передаточная функция полюсного ЦФ, соответствующего ЛРУ (4.12), определяется выражением

$$H(z) = \frac{Y(z)}{U(z)} = \frac{1}{1 + \sum_{m=1}^{M-1} a_m z^{-m}}.$$

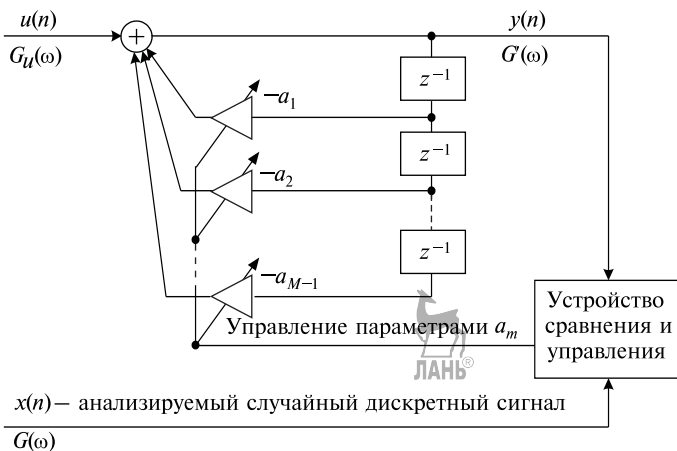


Рис. 4.11. Принцип параметрического цифрового спектрального оценивания на основе авторегрессионной модели дискретного случайного сигнала

Структура процедуры цифрового спектрального оценивания на основе АР-модели показана на рис. 4.11.

АЧХ ЦФ, представленного на рис. 4.11, имеет вид

$$A(\omega T) = \frac{1}{\sqrt{\left(1 + \sum_{m=1}^{M-1} a_m \cos m\omega T\right)^2 + \left(\sum_{m=1}^{M-1} a_m \sin m\omega T\right)^2}}.$$

По сравнению с выражением (4.11) определение спектральной плотности мощности сигнала $x(n)$ является более простым:

$$G(\omega) = \frac{\sigma^2 T}{\left(1 + \sum_{m=1}^{M-1} a_m \cos m\omega T\right)^2 + \left(\sum_{m=1}^{M-1} a_m \sin m\omega T\right)^2}.$$

Одним из этапов цифрового спектрального оценивания дискретных случайных процессов выступает процедура параметрической идентификации модели, заключающаяся в формировании оптимальных значений элементов множества $\{a_m\}$, обеспечивающих максимальную степень похожести модели и сигнала по используемому критерию. Способ идентификации выступает в качестве классификационного признака методов оценивания на основе АР-модели. Классификация методов параметрического цифрового спектрального оценивания на основе авторегрессионной модели случайного процесса, существенно дополняющая рис. 4.2, представлена на рис. 4.12.

Автокорреляционный метод расчета параметров модели авторегрессионного процесса. Ниже рассмотрен один из методов расчета параметров модели авторегрессионного процесса — автокорреляционный, который осуществляется на основе:

- блочной обработки N отсчетов анализируемого сигнала;
- метода линейного предсказания анализируемого сигнала;
- минимизации среднеквадратической ошибки прямого линейного предсказания.

Метод одношагового прямого линейного предсказания основан на использовании корреляционных связей между отсчетами обрабатываемого дискретного сигнала и выражается в предсказании текущего отсчета сигнала $x(nT)$ в виде линейной комбинации его предыдущих отсчетов:

$$x(nT) \approx \hat{x}(nT) = \sum_{m=1}^{M-1} a_m x(nT - mT) \quad \text{при } n = 0, 1, 2, \dots, N-1,$$

где $\hat{x}(nT)$ — предсказанное значение текущего отсчета; $M-1$ — порядок предсказания.

Для реализации автокорреляционного метода расчета параметров АР-модели выполняется цифровая фильтрация анализируемого сигнала $x(nT)$ нулевым фильтром с передаточной характеристикой $H_1(z)$:

$$H_1(z) = \frac{1}{H(z)} = \frac{E(z)}{X(z)} = 1 - \sum_{m=1}^{M-1} a_m z^{-m}, \quad (4.13)$$

что соответствует линейно-разностному уравнению

$$e(n) = x(n) - \sum_{m=1}^{M-1} a_m x(n-m) = x(n) - \hat{x}(n). \quad (4.14)$$

Дискретный сигнал $e(n)$, являющийся реакцией ЦФ, представляет собой ошибку линейного предсказания. Структурная схема фильтрации (4.14) показана на рис. 4.13.

Из уравнения (4.14) следует, что

$$x(n) = \sum_{m=1}^{M-1} a_m x(n-m) + e(n).$$

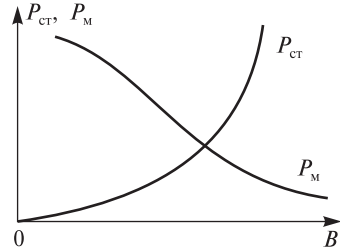


Рис. 4.12. Классификация методов цифрового спектрального оценивания на основе модели авторегрессионного процесса

вектор-столбец \mathbf{a} будет определяться из выражения $\mathbf{a} = \mathbf{R}\mathbf{B}^{-1}$, где \mathbf{B}^{-1} — матрица, обратная матрице \mathbf{B} . Матрица \mathbf{B} является теплицевой, т. е. симметрической и такой, что элементы главной диагонали и элементы, лежащие на параллелях к ней, соответственно равны между собой:

$$\mathbf{B} = \begin{bmatrix} B(0) & B(1) & B(2) & \dots & B(M-2) \\ B(1) & B(0) & B(1) & \dots & B(M-3) \\ B(2) & B(1) & B(0) & \dots & B(M-4) \\ \dots & \dots & \dots & \dots & \dots \\ B(M-2) & B(M-3) & B(M-4) & \dots & B(0) \end{bmatrix}.$$

Очевидно, что теплицевая матрица является невырожденной (ее определитель отличен от нуля), следовательно, система уравнений (4.31) является не только совместной, но и определенной, что означает единственность решения относительно \mathbf{a} .

Важным достоинством теплицевой матрицы \mathbf{B} является наличие эффективного итеративного алгоритма для ее обращения, получившего название алгоритма Левинсона—Дарбина. По этому алгоритму коэффициенты a_m , получаемые при решении системы (4.31) j -го порядка, рекуррентно выражаются через решение той же системы порядка $j-1$, что позволяет значительно снизить вычислительную сложность алгоритма идентификации. Используемые соотношения имеют вид:

$$\begin{cases} E_0 = B(0); \\ K_j = \frac{1}{E_{j-1}} \sum_{m=1}^{j-1} a_m^{(j-1)} B(j-m) - B(j); \\ a_m^{(j)} = \begin{cases} -K_j & \text{при } m = j; \\ a_m^{(j-1)} + K_j a_{j-m}^{(j-1)} & \text{при } 1 \leq m \leq j-1; \end{cases} \\ E_j = (1 - K_j^2) E_{j-1}, \end{cases} \quad (4.17)$$

где $a_m^{(j)}$ — коэффициенты анализирующего фильтра j -го порядка ($m = 1, 2, \dots, j$); K_j и E_j — промежуточные переменные, имеющие собственный физический смысл.

Система (4.17) решается последовательно для $j = 1, 2, \dots, M-1$, а окончанием процедуры идентификации является решение вида

$$a_m = a_m^{(M)}, \quad m = 1, 2, \dots, M-1.$$

Рассчитанные коэффициенты a_m делают возможным синтезировать фильтр (см. рис. 4.13). Значение дисперсии σ^2 дискретного белого шума $u(n)$, необходимого для создания АР-модели, опреде-

ляется следующим образом:

$$\sigma^2 = \frac{E}{N + M - 1}.$$

Известность параметров a_m и σ^2 позволяет выполнить расчет СПМ $G(\omega)$ анализируемого сигнала $x(nT)$, означающий, что задача по цифровому спектральному оцениванию решена.

Обязательным условием корректности анализа методом линейного предсказания является устойчивость АР-модели, для чего необходимо, чтобы все полюсы $H(z)$ лежали внутри единичного круга на z -плоскости с центром в начале координат. Одним из главных преимуществ автокорреляционного метода определения a_m является то, что обращение матрицы Тёплица гарантирует устойчивость используемой АР-модели.

Необходимо отметить, что цифровое спектральное оценивание, использующее авторегрессионную модель, в настоящее время широко применяется в устройствах параметрического кодирования речевого сигнала на основе метода линейного предсказания речи (липредерах). При этом основным отличием от изложенных выше принципов является то, что в липредерах в качестве сигнала возбуждения адаптивного ЦФ выступает, как правило, не белый шум, а специально сформированный сигнал, адекватно отображающий генераторную функцию артикуляционного аппарата человека.

Очевидно, что использование методов на основе АР-модели желательно для спектрального оценивания таких сигналов, которые по своим свойствам близки к авторегрессионным.

Для сравнительного анализа параметрического спектрального оценивания методами на основе модели авторегрессионного процесса при блочной обработке отсчетов использован тестовый дискретный сигнал $\text{test}(n)$. Эффективность различных методов параметрического спектрального оценивания в этом случае может быть визуально оценена по способности обнаружения в спектре 256-точечного зашумленного сигнала гармонических составляющих $\cos(0,275\pi n)$ и $\cos(0,2\pi n)$, расположенных на нормированных частотах $f_{01} = 0,1375$ и $f_{02} = 0,1$ соответственно.

На рис. 4.14 представлены односторонние спектральные плотности мощности сигнала $\text{test}(n)$, показанные в децибелах, для диапазона нормированной частоты $0 \leq f_0 \leq 0,5$. Как правило, точность спектрального оценивания возрастает при увеличении порядка АР-модели. С другой стороны, при одинаковом порядке АР-моделей рассматриваемая группа методов часто дает очень близкие результаты спектрального оценивания. Так, из СПМ, представленных на

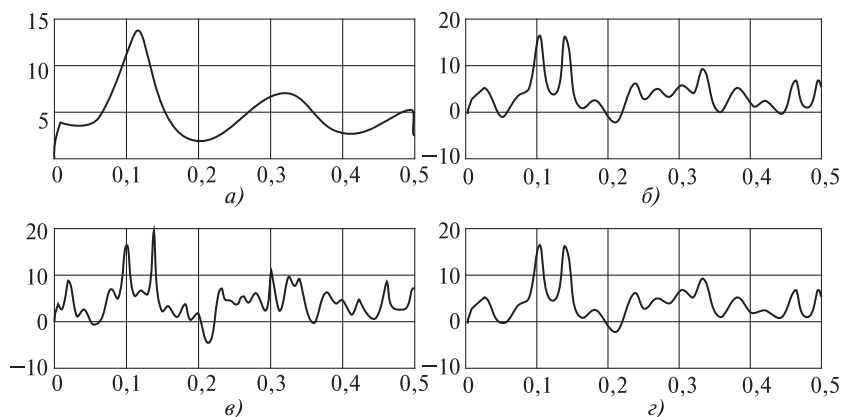


Рис. 4.14. Спектральные плотности мощности сигнала $\text{test}(n)$, полученные методом Берга на основе АР-моделей 10-го (а) и 50-го (б) порядков, модифицированным ковариационным методом (в) и методом Юла—Уолкера (з) на основе АР-моделей 30-го порядка

рис. 4.14, лучшее частотное разрешение обеспечивает метод Берга при использовании модели 50-го порядка (рис. 4.14,а), а худшее, не позволяющее различить гармонические компоненты, — этот же метод при использовании модели 10-го порядка (рис. 4.14,б).

Не менее показательным является очевидная близость оценок СПМ, полученных модифицированным ковариационным методом (рис. 4.14,в) и методом Юла—Уолкера (рис. 4.14,з) при одинаковом (30-м) порядке моделей.

4.4.4. Параметрическое спектральное оценивание на основе модели процесса скользящего среднего

Модель процесса скользящего среднего (МА — Moving Average) определяется частным случаем выражения (4.10) при отсутствии рекурсивной части ($a_0 = 1$, $a_m = 0$, $m = 1, 2, 3, \dots$) и имеет вид

$$y(nT) = \sum_{k=0}^{K-1} b_k u(nT - kT).$$

Следовательно, в методах параметрического цифрового спектрального оценивания на основе СС-модели в качестве формирователя модели анализируемого сигнала выступает адаптивный НРЦФ.

Структура процедуры цифрового спектрального оценивания на основе модели процесса скользящего среднего показана на рис. 4.15.

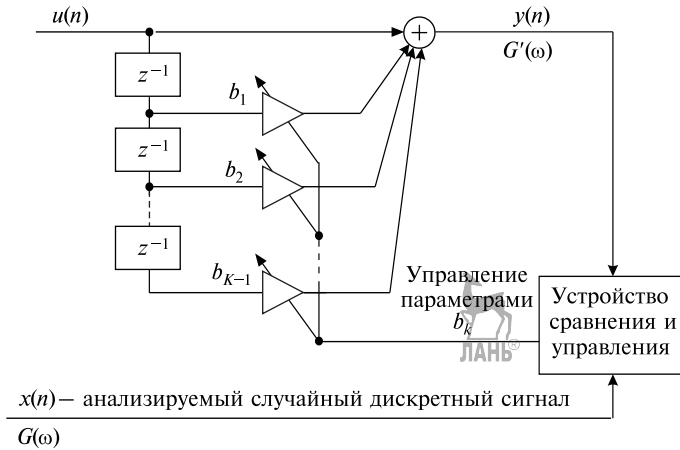


Рис. 4.15. Принцип параметрического цифрового спектрального оценивания на основе модели процесса скользящего среднего

Амплитудно-частотная характеристика НРЦФ имеет вид

$$A(\omega T) = \sqrt{\left(\sum_{k=0}^{K-1} b_k \cos k\omega T\right)^2 + \left(\sum_{k=0}^{K-1} b_k \sin k\omega T\right)^2}.$$

При условии приближенного равенства сигналов $x(n)$ и $y(n)$ искомая спектральная плотность мощности анализируемого сигнала выражается через дисперсию шумового воздействия и параметры цифрового фильтра:

$$G(\omega) = \sigma^2 T \left[\left(\sum_{k=0}^{K-1} b_k \cos k\omega T\right)^2 + \left(\sum_{k=0}^{K-1} b_k \sin k\omega T\right)^2 \right].$$

Очевидно, что точность аппроксимации сигнала $x(nT)$ в методах на основе СС-модели определяется порядком используемого цифрового фильтра $(K - 1)$. Преимуществом данных методов является гарантированная устойчивость нерекурсивного адаптивного цифрового фильтра.

4.4.5. Параметрическое спектральное оценивание на основе модели процесса авторегрессии — скользящего среднего

Параметрическая модель процесса авторегрессии — скользящего среднего (ARMA — Autoregressive Moving Average) является обобщающей для методов параметрического цифрового спектрального оценивания на основе представления анализируемых сигналов в виде реакции линейных цифровых фильтров на дискретный белый шум.

В методах спектрального оценивания на основе данной модели в качестве формирователя модели анализируемого сигнала выступает адаптивный полюсно-нулевой цифровой фильтр. Структура процедуры цифрового спектрального оценивания на основе модели АРСС-процесса показана на рис. 4.16.

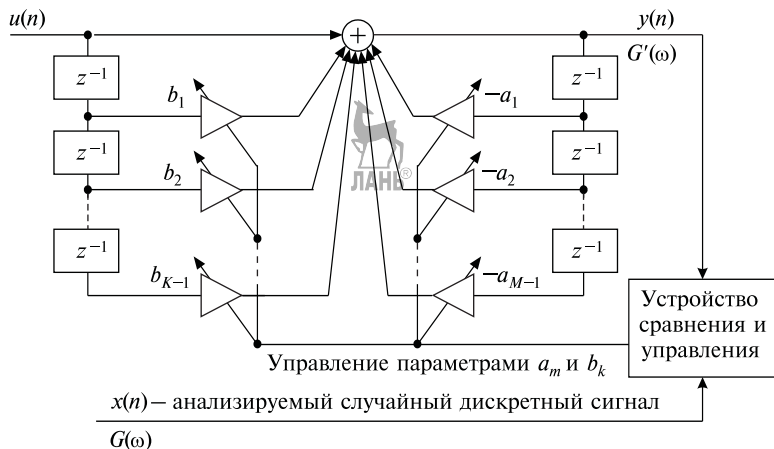


Рис. 4.16. Принцип параметрического цифрового спектрального оценивания на основе модели процесса авторегрессии — скользящего среднего

Точность аппроксимации сигнала $x(nT)$ в методах на основе модели процесса АРСС определяется порядком как нерекурсивной части используемого цифрового фильтра $(K - 1)$, так и его рекурсивной части $(M - 1)$.



5 Цифровая обработка речевого сигнала



5.1. Процесс речеобразования

Процесс речеобразования начинается с того момента, когда мысленное абстрактное сообщение должно достичь собеседника при помощи артикуляционного аппарата. Сообщение преобразуется в последовательность нервных импульсов, управляющих элементами артикуляции (голосовые связки, язык, зубы, губы и т.д.), которые формируют акустическое речевое колебание, несущее информацию об исходном сообщении.

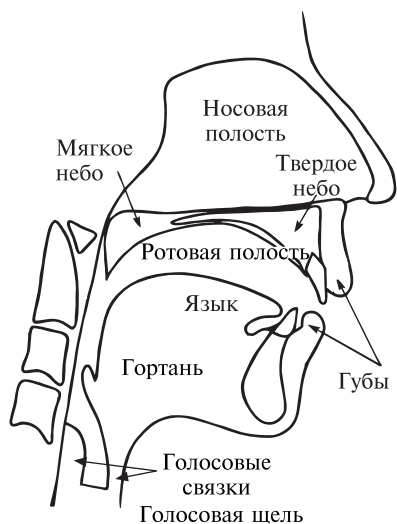


Рис. 5.1. Речевой аппарат человека

Артикуляционный аппарат человека наглядно демонстрируется на рис. 5.1.

Голосовой тракт начинается с прохода между голосовыми связками, называемого *голосовой щелью*, и заканчивается у губ. Голосовой тракт состоит из *гортани* и рта, или *ротовой полости*. *Носовая полость* начинается у небной занавески (мягкого неба) и заканчивается ноздрями. При опущенной небной занавеске носовая полость акустически соединена с голосовым трактом и участвует в образовании носовых звуков речи. Легкие, бронхи и трахея, расположенные ниже голосовых связок,

хотя и не входят в голосовой тракт, но участвуют в создании воздушного потока, являющегося исходным сигналом для дальнейшего преобразования в речевое колебание. В совокупности эти органы формируют сигнал возбуждения голосового тракта и, таким об-

разом, являются источником энергии для образования речи. Речь представляет собой акустическую волну, которая вначале излучается этой системой при выталкивании воздуха давлением, создаваемым легкими, и затем преобразуется в голосовом тракте: многократно отражается от его стенок, вызывает их вибрацию, после чего излучается через нос и рот.

Голосовой тракт и носовая полость представлены на рис. 5.2 в виде труб с переменной площадью поперечного сечения.

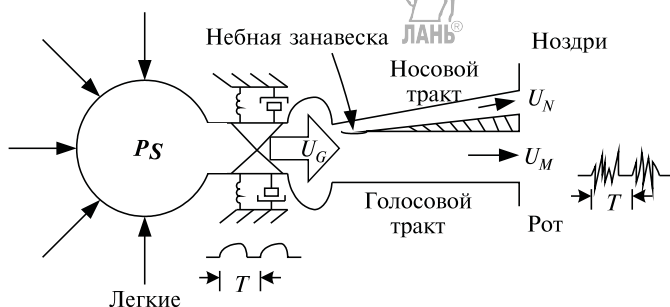


Рис. 5.2. Физическая модель механизма речеобразования: P_S — давление, создаваемое легкими при выталкивании воздуха; U_G — сигнал, формируемый голосовой щелью из воздушного потока, создаваемого легкими; T — период смыкания, размыкания голосовых связок, моделируемых резонансными контурами; U_N — акустический сигнал, излучаемый через ноздри; U_M — акустический сигнал, излучаемый через губы, зубы

При прохождении звуковых волн через эти трубы их частотный спектр изменяется в соответствии с частотной избирательностью трубы. Этот эффект похож на резонансные явления в музыкальных инструментах. При описании речеобразования резонансные частоты трубы голосового тракта называются *формантными частотами* или *формантами*. Формантные частоты зависят от размеров и формы голосового тракта. Различные звуки образуются путем изменения формы голосового тракта. Спектральные свойства речевого сигнала изменяются во времени в соответствии с изменением формы голосового тракта.

Все звуки речи могут быть разделены на три четко выраженные группы по виду сигнала возбуждения голосового тракта: вокализованные, невокализованные и взрывные.

Вокализованные звуки (гласные, звонкие согласные: а, о, у, з, ...) образуются проталкиванием воздуха через голосовую щель, при этом периодически, с периодом T , напрягаются и расслабляются голосовые связки — таким образом возникает квазипериодическая последовательность импульсов потока воздуха, возбуждающая

голосовой тракт. В результате акустическое колебание, излучаемое через ноздри, губы и зубы, представляет собой квазипериодический сигнал. В модели речеобразования на рис. 5.2 периодическое сокращение голосовых связок моделируется двумя резонансными контурами. Их совместное колебание управляет площадью отверстия, формирующего сигнал возбуждения.

Невокализованные звуки (глухие согласные: ϕ , x , $ш$, ...) образуются при сужении голосового тракта в каком-либо месте и проталкивании воздуха через суженное место с высокой скоростью, достаточной для образования шумового воздушного потока. После этого шумовой сигнал возбуждения преобразуется в голосовом тракте — происходит перераспределение энергии по частотному диапазону в соответствии с частотной избирательностью голосового тракта.

При произношении *взрывных* звуков ($б$, $п$, ...) голосовой тракт полностью закрывается, обычно в начале. За этим местом возникает повышенное сжатие воздуха. Затем воздух резко высвобождается и формируется акустическое колебание, излучаемое голосовым трактом.

Многообразие звуков речи на ограничивается указанными тремя группами, но этих групп достаточно для построения моделей речеобразования, полученных на основе акустической теории речи.

Наибольшее распространение получила модель речеобразования (рис. 5.3), учитывающая различную природу сигналов возбуждения, характер прохождения акустической волны через голосовой тракт и излучения речевого колебания.

Обобщенная модель имеет ряд ограничений, связанных с характером изменения параметров во времени, необходимостью измене-

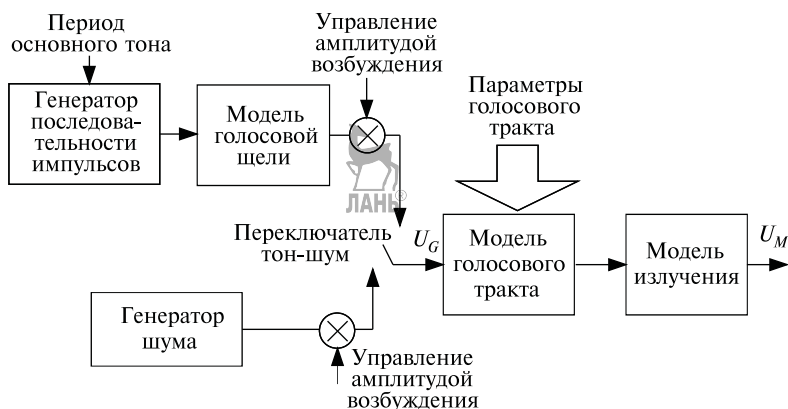


Рис. 5.3. Обобщенная дискретная модель речеобразования

ния модели для точного описания носовых звуков, упрощенным разделением типов возбуждения на вокализованное и невокализованное. В случае необходимости представленная модель речеобразования может быть изменена для корректного описания отдельных видов звуков. Важной задачей при исследовании динамики изменения характеристик РС является выбор длительности его временного сегмента, на которые он разделяется при обработке. С одной стороны, он должен быть достаточно мал, чтобы последовательность сегментов более точно отражала кратковременную динамику изменения РС, а с другой стороны, достаточно большим, чтобы последовательность сегментов более точно отражала долговременную динамику РС. Также сегменты РС могут разделяться по признаку соответствия фонетическим образам (фонема, слог, слово, команда), по признаку активности речи (паузы, сегменты активной речи), активная речь в свою очередь может разделяться на вокализованные и невокализованные сегменты.



5.2. Характеристики речевого сигнала

Речевым сигналом (РС) называется электрическое колебание, наблюдаемое на выходе акустико-электрического преобразователя при воздействии на его вход акустической речевой волны. К настоящему времени подробно описаны все существенные характеристики речевого сигнала и сформирован широкий набор математических моделей его представления: от шумовых случайных на невокализованных участках речи до периодических детерминированных на вокализованных участках. Применительно к задачам электросвязи общепринятым является вероятностное описание РС. Смысловое содержание речевой волны априори неизвестно, и, кроме того, одному смысловому содержанию в разных экспериментах могут соответствовать различные электрические колебания, отличающиеся расположением во времени, формой, длительностью, т. е. РС в общем случае можно рассматривать как непрерывный по времени и состоянию случайный процесс, реализации которого наблюдаются на выходе акустико-электрического преобразователя.

По своей природе РС нестационарен, что составляет одну из основных сложностей его адекватного математического описания и моделирования. Процедуры обработки нестационарных процессов представляют собой сложную, а с точки зрения оптимальности — часто неразрешимую задачу. Поэтому во многих случаях является оправданным переход к более простым стационарным моделям.

Стремление учесть нестационарную природу РС и относительная простота анализа и обработки стационарных сигналов обусло-

вили широкое распространение локально-стационарной модели речевого сигнала $S(t)$ в виде реакции динамической системы с сосредоточенными параметрами, реализующей оператор Π , на соответствующее случайное воздействие $\xi(t)$:

$$S(t) = \Pi[\xi(t)] \quad (5.1)$$

Очевидно, что модель (5.1) может быть представлена и для дискретной временной шкалы. Одним из преимуществ рассматриваемой модели является возможность математического описания РС с помощью положений теорий динамических систем и случайных процессов. Кроме того, рассматриваемая модель является достаточно хорошей аппроксимацией объекта, порождающего речевой сигнал, — артикуляционного аппарата: оператор Π представляет собой эквивалентную модель голосового тракта, а случайный сигнал возбуждения $\xi(t)$ — исходный турбулентный шум.

Временная длительность участка локальной стационарности определяется интервалом квазистационарности речевого сигнала, составляющим от 10 до 30 мс.

Вероятностная модель РС определяется совокупностью вероятностных характеристик случайного процесса. На практике РС описывается статистическими характеристиками, полученными по опытным данным конечного объема.

Основными статистическими характеристиками РС являются:

- амплитудное распределение;
- спектральное распределение;
- распределение формант по частоте и амплитуде;
- распределение частоты основного тона;
- распределение длительностей речевых участков и пауз;
- звуковой состав русской речи.

Амплитудное распределение русской речи показано на рис. 5.4. Эта характеристика представляет собой интегральное распределение относительно числа интервалов времени наблюдения N , в которых превышает тот или иной уровень речи B , характеризующий интенсивность речи во всем частотном диапазоне.

Средний уровень речи B (условно обозначенный нулем) превышает в 19 % интервалов времени наблюдения. За максимальный уровень принимают уровень речи, который превышает лишь в 1 % интервалов времени наблюдения, — это уровень, равный 12 дБ; за минимальный уровень — который превышает в 99 % интервалов времени наблюдения, — это уровень, равный —35 дБ.

Динамический диапазон D РС определяется как разность между максимальным и минимальным уровнями и равен 47 дБ. Пик-

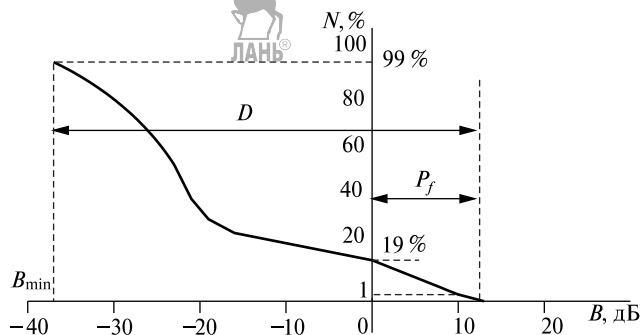


Рис. 5.4. Амплитудное распределение русской речи

фактор P_f — это разность между максимальным и средним уровнями речи, равная 12 дБ.

Спектр речи $B_p(f)$ есть ее СПМ в течение длительного времени наблюдения. Значение СПМ речи на конкретной частоте f определяется выражением вида

$$B_p(f) = 20 \lg(P_1/P_2),$$

где P_1 — звуковое давление на частоте f , отнесенное к полосе шириной 1 Гц, Па; P_2 — звуковое давление на частоте f , соответствующее абсолютно нулевому уровню ($2 \cdot 10^{-5}$ Па).

Спектр русской речи, усредненный для мужских и женских голосов, представлен на рис. 5.5.

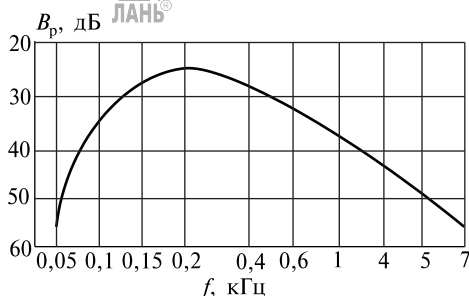


Рис. 5.5. Спектр русской речи

Каждому звуку речи соответствует свое распределение энергии по частотному диапазону, называемое формантным рисунком. Области частотного диапазона, где происходит увеличение амплитуды спектральных составляющих, называются *формантными областями*, а частоты, на которых происходит максимальное увеличение амплитуды, — *формантными частотами*.

Форманта характеризуется амплитудой A_i , частотой F_i и шириной полосы ΔF_i , или добротностью Q_i . Ширина i -й форманты ΔF_i определяется на уровне $0,707 A_i$ и связана для гласных звуков с добротностью соотношением $\Delta F_i = F_i / Q_i$.

Различные звуки имеют разное число формант: гласные — до четырех, глухие согласные до 5–6 формант. Первые две форманты называются *основными*, остальные — *вспомогательными*. Основные форманты определяют произносимый звук речи, а вспомогательные характеризуют индивидуальную для каждого человека окраску, тембр речи.

Важной характеристикой вокализованных звуков является *частота основного тона* — частота колебаний голосовых связок или частота первой гармоники спектра вокализованных звуков. *Период основного тона* $T_{от}$ — величина, обратная частоте основного тона. У вокализованных звуков спектр является дискретным, и все спектральные составляющие имеют частоту, кратную частоте основного тона. Частота основного тона изменяется в пределах от 60...70 Гц для мужских голосов до 450...500 Гц для высоких женских голосов. Средняя частота основного тона для мужских голосов 130...150 Гц, для женских — 250 Гц.

При разработке систем кодирования речи важную роль играет *распределение длительностей речевых участков и пауз*. Длительность отдельных звуков (фонем) речи составляет 20...350 мс. При этом гласные звуки имеют большую длительность, чем согласные. Так как максимальная длительность звука равна 350 мс, паузой считается отсутствие речи в течение времени, большего 350 мс. Средняя длительность пауз составляет приблизительно 16...20 % длительности речи, а средняя скорость речи 10–15 звуков/с. Темп речи может изменяться в широких пределах. Длительность звуков и пауз также изменяется, причем длительность гласных звуков изменяется в большей степени. Относительная длительность согласных тем больше, чем выше темп речи: при очень быстром темпе отдельные гласные могут полностью исчезать, при больших замедлениях темпа длительность согласных звуков практически не изменяется, а удлинение слогов происходит за счет гласных звуков.

Хотя звуки существенно отличаются друг от друга, имеется ряд особенностей, характерных для отдельных групп фонем. Например, для вокализованных можно выделить несколько участков сигнала: начальный (приблизительно 0,04 с) — рост амплитуды сигнала; средний участок (приблизительно 0,165 с) — многократное повторение практически одинакового сигнала; конечный участок (приблизительно 0,09 с) — уменьшение амплитуды сигнала.

Таблица 5.1

Основные статистические характеристики и параметры речевого сигнала

Характеристика	Производные параметры	Естествен- ный РС	Телефон- ный РС
Амплитудное распределение	Мощность в точке с нулевым измеритель- ным уровнем на интервале активности речи, мкВт: максимальная средняя минимальная Динамический диапазон, дБ Пик-фактор, дБ	5000 315 0,01 57 12	2200 88 0,1 43 14
Спектральное распределение	Полоса частот, Гц: минимальная частота максимальная частота	50...100 8000...15000	300 3400
Распределение формант по час- тоте и амплиту- де	Диапазон частот для формант вокализо- ванных звуков, Гц: первая форманта вторая форманта третья форманта Диапазон частот для формант невокали- зованных звуков, кГц: первая форманта вторая форманта антиформанта Динамический диапазон формант, дБ	150...900 550...2800 1500...3400 1...3,5 2,5...6 1,5...4 25	1...3,4 2,5...3,4 1,5...3,4
Распределение частоты основ- ного тона	Частота основного тона для мужских го- лосов, Гц: диапазон значений среднее значение среднеквадратическое отклонение Частота основного тона для женских го- лосов, Гц: диапазон значений среднее значение среднеквадратическое отклонение Максимальная скорость изменения частоты основного тона, Гц/с	56...240 128 35 140...450 256 50 6000	
Распределение длительностей речевых участ- ков и пауз	Коэффициент активности речи Средняя длительность активных сегмен- тов речи, с Средняя длительность пауз, с Частота следования пауз, Гц	0,25...0,35 0,35 0,184 2	
Звуковой состав русской речи	Средняя длительность звуков, с: вокализованных невокализованных Вероятность появления на интервале ак- тивности речи: вокализованных звуков невокализованных звуков Количество фонем Средний темп речи, фонем/с	0,12 0,06 0,77 0,23 42 10...12	

Длительность отдельных участков неодинакова для звуков, произносимых в различных сочетаниях. У взрывных звуков практически нет среднего участка, только нарастание и спад.

При анализе звукового состава русской речи звуки разделяются на гласные и согласные, которые в свою очередь разделяются на ударные, безударные, твердые, мягкие, звонкие, глухие и проч.

Гласные составляют примерно 43,5 %, согласные — 56,5 % общего числа звуков. При этом невокализованные звуки составляют примерно 32 %. Наиболее распространенный гласный звук — это *а*, самый распространенный согласный — *т*. Среди гласных звуков наиболее редким является звук *э*, а среди согласных — *фь*.

Оценки указанных характеристик разными исследователями имеют некоторые отличия, что объясняется различными объемами исследованного РС, использованными методами измерений, длительностями единичных исследуемых сегментов речи.

Естественный речевой сигнал по своей природе обладает как статистической, так и психофизической избыточностью. Наиболее яркими проявлениями статистической избыточности являются наличие пауз между элементами речи и квазипериодичность вокализованных сегментов РС. Психофизическая избыточность выражается в нечувствительности потребителя к некоторым необратимым искажающим преобразованиям РС при восприятии его органами слуха. Избыточность речи позволяет ввести в рассмотрение телефонный РС, параметры которого нормированы Международным союзом электросвязи (МСЭ), имеют уменьшенные диапазоны значений, но обеспечивают достаточное качество телефонии. В табл. 5.1 представлены обобщенные параметры РС, производные от перечисленных характеристик, для естественного и телефонного РС.

5.3. Метод линейного предсказания дискретного речевого сигнала

Линейное предсказание (ЛП) является одним из наиболее эффективных методов анализа РС, который становится доминирующим при оценке основных параметров речи, таких как период основного тона, форманты, спектр и др., а так же при ее компрессии с целью низкоскоростной передачи и экономного хранения.

Основные положения метода линейного предсказания хорошо согласуются с моделью речеобразования (рис. 5.3), где показано, что речевой сигнал можно представить в виде сигнала на выходе линейной системы с переменными во времени параметрами, возбуждаемой квазипериодическими импульсами (в пределах вокализованного сегмента) или случайным шумом (на невокализованном сегменте).

Отличительным признаком метода ЛП является одношаговое приближенное представление (предсказание) текущего отсчета $s(n)$ РС линейной комбинацией $s_{\text{пр}}(n)$ конечного числа предыдущих отсчетов:

$$s(n) \approx s_{\text{пр}}(n) = \sum_{m=1}^M a_m s(n-m), \quad (5.2)$$

где n — номер текущего отсчета; m — задержка, выраженная числом периодов дискретизации; M — порядок предсказания; a_m — коэффициенты предсказания, $m = 1, 2, \dots, M$.

Коэффициенты предсказания при этом определяются однозначно минимизацией среднего квадрата разности между отсчетами речевого сигнала и их предсказанными значениями (на конечном интервале).

Следует отметить, что последовательность, удовлетворяющая линейно-разностному уравнению (5.2), является авторегрессионной, так как $s(n)$ совпадает с условным математическим ожиданием

$$E \left\{ \frac{s(n)}{s(n-1), s(n-2), \dots, s(n-M)} \right\}$$

или, иначе говоря, с регрессией $s(n)$ по $s(n-1), s(n-2), \dots, s(n-M)$.

Возможность линейного прогнозирования по формуле (5.2) текущего отсчета речевого сигнала (рис. 5.6) объясняется:

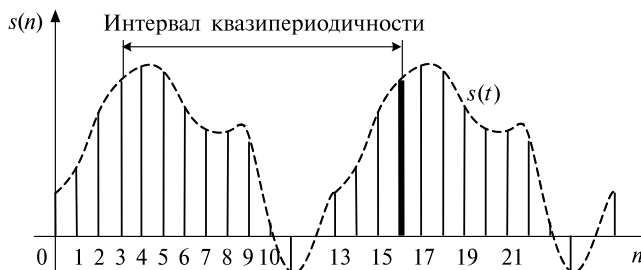


Рис. 5.6. Иллюстрация наличия корреляционных связей между отсчетами дискретного речевого сигнала

- наличием корреляционных связей между соседними отсчетами РС при его равномерной дискретизации с частотой, выбираемой из расчета максимально возможной ширины спектра источника, что обусловлено неполнотой априорных сведений об обрабатываемом РС; соответствующее линейное предсказание часто называют *кратковременным*;

- наличием корреляционных связей между отсчетами РС, разделенными интервалами, близкими к периоду основного тона (интервалами квазипериодичности), — такое линейное предсказание часто называют *долговременным*.

Переход от формулы (5.2) к точному равенству требует учета сигнала ошибки предсказания $c(n)$:

$$s(n) = s_{\text{пр}}(n) + c(n) = \sum_{m=1}^M a_m s(n-m) + c(n). \quad (5.3)$$

Повышение порядка предсказания M приводит к повышению точности процедуры прогнозирования (уменьшению ошибки предсказания текущего отсчета, уменьшению мощности сигнала ошибки предсказания $c(n)$) и в предельном случае ($M \rightarrow \infty$) обеспечивает устранение сигнала ошибки предсказания ($c(n) \rightarrow 0$), что означает равенство значения текущего отсчета его предсказанному значению: $s(n) = s_{\text{пр}}(n)$.

На рис. 5.6 пунктиром для наглядности показан аналоговый РС $s(t)$, дискретизацией которого получен дискретный РС $s(n)$. Очевидно, что, например, выделенный на рисунке отсчет $s(16)$ в значительной степени определяется как предыдущими отсчетами $s(15)$, $s(14)$, $s(13)$, $s(12)$, так и отсчетами $s(3)$, $s(2)$, $s(1)$, удаленными от него на интервалы времени, близкие к интервалу квазипериодичности. Кратковременное ЛПП при конечном порядке предсказания более эффективно на участках монотонного изменения РС, а долговременное ЛПП в большей степени проявляется на вокализованных (квазипериодических) участках РС.

Практическая реализация метода ЛПП требует конечного значения M , определяемого допустимой погрешностью между речевым сигналом и его синтезированной копией. В наиболее популярных алгоритмах липредерного кодирования значение порядка предсказания M варьируется в пределах от 8 (например, в стандарте GSM) до 50 (в стандарте G.728).

Очевидно, что в качестве синтезирующего фильтра линейного предсказания, реализующего уравнение (5.3), должен использоваться рекурсивный цифровой фильтр M -го порядка, структурная схема которого показана на рис. 5.7, а передаточная функция определяется следующим выражением:

$$H(z) = \frac{S(z)}{C(z)} = \frac{1}{1 - \sum_{i=1}^M a_i z^{-i}}. \quad (5.4)$$

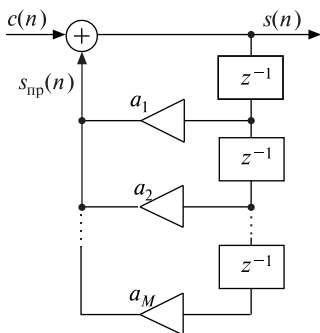


Рис. 5.7. Структурная схема синтезирующего фильтра линейного предсказания

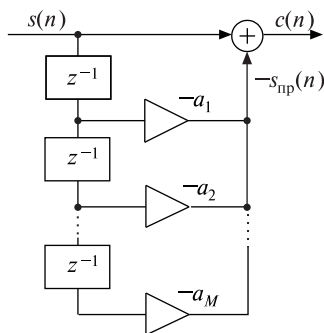


Рис. 5.8. Структурная схема анализирующего фильтра линейного предсказания

Передаточная функция $A(z)$ анализирующего фильтра ЛП, обратная по отношению к $H(z)$:

$$A(z) = \frac{1}{H(z)} = \frac{C(z)}{S(z)} = 1 - \sum_{i=1}^M a_m z^{-m}. \quad (5.5)$$

Из формулы (5.5) следует линейно-разностное уравнение анализирующего фильтра ЛП:

$$c(n) = s(n) - \sum_{m=1}^M a_m s(n-m) = s(n) - s_{\text{пр}}(n),$$

которое реализуется нерекурсивным фильтром M -го порядка.

Структурная схема такого фильтра представлена на рис. 5.8.

Локальная стационарность анализирующего и синтезирующего фильтров ЛП на длительности сегмента РС относит задачу оценки параметров $\{a_m\}$ по наблюдаемой реализации $s(n)$ к классу задач параметрической идентификации авторегрессионных стационарных систем во временной области, которая была рассмотрена в главе 4. Очевидно, что в качестве настраиваемой модели выступает рекурсивный цифровой фильтр (5.4), синтезирующий речевой сигнал.

Критерием идентификации, как правило, выступает минимум среднеквадратической ошибки (СКО) предсказания, выражающийся для речевого сегмента, содержащего N отсчетов, в виде

$$E_{\text{пр}} = \frac{1}{N} \sum_{n=0}^{N-1} [s(n) - s_{\text{пр}}(n)]^2. \quad (5.6)$$

Совокупность параметров a_m , минимизирующих (5.6), соответ-

ствуется решению системы линейных уравнений

$$\frac{\partial E}{\partial a_m} = 0 \quad \text{при } m = 1, 2, 3, \dots, M-1,$$

которая выражается в аналитической форме через значения автокорреляционной функции $B(\tau)$, $\tau = 0, 1, 2, \dots, M$, исходного РС:

$$\begin{aligned} B(0)a_1 + B(1)a_2 + B(2)a_3 + \dots + B(M-1)a_M &= B(1); \\ B(1)a_1 + B(0)a_2 + B(1)a_3 + \dots + B(M-2)a_M &= B(2); \\ B(2)a_1 + B(1)a_2 + B(0)a_3 + \dots + B(M-3)a_M &= B(3); \\ &\dots\dots\dots \\ B(M-1)a_1 + B(M-2)a_2 + B(M-3)a_3 + \dots + B(0)a_M &= B(M). \end{aligned} \quad (5.7)$$

Система уравнений (5.7) известна как система Юла-Уокера и в матричной форме может быть записана в виде $\mathbf{B}\mathbf{a} = \mathbf{R}$. Очевидно, что искомым вектор-столбец \mathbf{a} будет определяться из выражения $\mathbf{a} = \mathbf{B}^{-1}\mathbf{R}$, где \mathbf{B}^{-1} — матрица, обратная матрице \mathbf{B} . Матрица

$$\mathbf{B} = \begin{bmatrix} B(0) & B(1) & B(2) & \dots & B(M-1) \\ B(1) & B(0) & B(1) & \dots & B(M-2) \\ B(2) & B(1) & B(0) & \dots & B(M-3) \\ \dots & \dots & \dots & \dots & \dots \\ B(M-1) & B(M-2) & B(M-3) & \dots & B(0) \end{bmatrix}$$

является тѐплицевой, т.е. симметрической и такой, что элементы главной диагонали и элементы, лежащие на параллелях к ней, соответственно равны между собой. Существенно, что тѐплицевая матрица является невырожденной (ее определитель отличен от нуля), а следовательно, система уравнений (5.7) является не только совместной, но и определенной, что означает единственность решения относительно \mathbf{a} .

Обязательным условием корректности анализа методом линейного предсказания является устойчивость синтезирующей системы $H(z)$, для чего необходимо, чтобы все полюсы лежали внутри единичного круга на z -плоскости с центром в начале координат. Полюсы системы $H(z)$ определяются нулями многочлена (5.5). Одним из главных преимуществ изложенного метода определения $\{a_m\}$, часто называемого корреляционным, является то, что обращение матрицы Теплица гарантирует устойчивость системы (5.7), что обеспечивает синтез минимально-фазового полюсного фильтра. Таким образом, обеспечивается затухание (стремление к нулю) реакции синтезирующего цифрового фильтра на любое ограниченное воздействие.

Вследствие тесной связи линейного предсказания со спектральным анализом коэффициенты предсказания не только являются ре-

зультатом идентификации синтезирующего фильтра ЛП в временной области, но и несут в себе комбинированную информацию о спектральной огибающей анализируемого сегмента речи. Сглаженный характер получаемого амплитудного спектра (рис. 5.9) иллюстрирует способность анализа методом линейного предсказания отделять сигнал возбуждения от характеристик голосового тракта, что полностью соответствует представленной модели речеобразования.

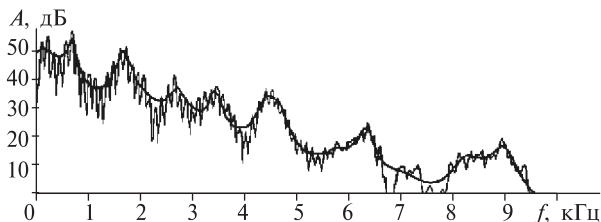


Рис. 5.9. Амплитудный спектр РС, полученный методом ЛП при $M = 28$ (изображен на фоне преобразования Фурье)

С учетом того, что в частотной области спектр коротких отрезков РС можно представить в виде произведения огибающей, характеризующей состояние голосового тракта, и сигнала возбуждения, указанное свойство метода линейного предсказания создает условия для значительного снижения скорости кодирования речевого источника по сравнению с методами, использующими кратковременный спектральный анализ на основе преобразования Фурье.

5.4. Эффективное кодирование речевого сигнала

5.4.1. Задачи, параметры и классификация кодеков речи

Постоянный рост объемов информации, циркулирующей в современных телекоммуникационных системах, сопровождается сохранением доминирующей роли речевых сообщений. Известно, что в большинстве корпоративных ИТКС телефонный трафик является основным и составляет более половины общего объема трафика. Речь, являясь естественной опорой коммуникации и наиболее универсальным инструментом человеческого общения, не только сохраняет свою значимость в период активной эволюции человеческого общества в состояние информационного, но и приобретает еще большую важность, обусловленную постоянным возникновением новых видов сервиса, основанных на синтезе эксклюзивных свойств речи и растущих возможностей современных телекоммуникационных технологий.

Речевой кодек (РПУ, речепреобразующее устройство) — это устройство преобразования РС в цифровую форму на передающей стороне и восстановления его аналоговой формы на приемной стороне. Следует отметить, что речевой кодек является одним из важнейших элементов современных ИТКС при передаче по ним речевой информации, так как качественные показатели процедуры кодирования РС определяют верхнюю границу качества передачи речи в ИТКС на основе цифровых телекоммуникационных систем.

Основными задачами синтеза речевых кодеков традиционно являются повышение качества речепреобразования при сохранении скорости кодирования РС или снижение скорости кодирования без ухудшения качественных оценок синтезированной речи.

Учет задач речевого кодирования и тенденций развития телекоммуникационных систем позволяет сформировать комплекс требований к перспективным речевым кодекам. К числу основных требований следует отнести:

- высокое качество синтезированной речи;
- низкую среднюю скорость кодирования речи;
- малую алгоритмическую задержку речевого сигнала в процессе кодирования (задержка передачи РС от одного абонента другому, одной из составляющих которой является алгоритмическая задержка в процессе кодирования (декодирования), вызывает два нежелательных явления — появление эхо-сигналов, когда абонент прослушивает свой голос, и наложение речи — процесс, при котором речь одного абонента прослушивается другим абонентом в тот момент, когда он ведет активный разговор.

Влияние задержки на качество звучания синтезированного РС рассмотрено в рекомендации ITU-T G.114 Annex B. В указанной рекомендации определены следующие границы значений задержки:

- 1) 0...150 мс — задержка допустима для большинства приложений;
- 2) 150...400 мс — задержка допустима, когда абонент осведомлен о ее существенном влиянии на качество предоставляемой услуги;
- 3) более 400 мс — недопустимая задержка;

Каждое из указанных ниже требований формирует вектор желательного изменения соответствующего параметра РПУ, комплекс которых во многом соответствует комплексу параметров кодирования РС:

- высокая устойчивость к канальным ошибкам;
- возможность работы в широком динамическом диапазоне входных воздействий сигналов и шумов;

- незначительное снижение качества речи при каскадном соединении кодеков;
- хорошее воспроизведение неречевых звуков, издаваемых абонентами при телефонном разговоре, а также фоновых звуков;
- простота сопряжения со смежными устройствами мультимедийных абонентских терминалов;
- низкая вычислительная сложность алгоритмов;
- низкая требуемая емкость устройств памяти;
- соответствие элементной базе и тенденциям ее развития;
- малое энергопотребление и низкую стоимость;
- низкие массогабаритные показатели.

Представленная совокупность требований характеризует общий подход к разработке и совершенствованию речепреобразующих устройств. В качестве критичных требований для решения конкретных задач может выступать любая совокупность требований из представленного перечня, а также иные требования.

Процедура речевого кодирования, реализуя перевод РС из непрерывного по времени и состоянию вида в цифровой, одновременно, как правило, решает и задачу сжатия речи. Задача сжатия речи является частным случаем задачи эффективного кодирования РС в условиях идеального канала и в общем случае представляет собой задачу его компактного дискретного представления в условиях недостаточной априорной информации о свойствах источника за счет сокращения или даже полного устранения естественной статистической избыточности РС (статистическое кодирование источника) и (или) использования уменьшающих исходный объем данных необратимых преобразований, которые не снижают субъективного качества восстановленной речи, т. е. сокращения психофизической избыточности (психофизическое кодирование источника).

Очевидно, что эффективность различных методов кодирования РС может оцениваться степенью достижимого сжатия сигнала при условии возможности восстановления на приемной стороне исходного сообщения с требуемым качеством в соответствии с принятым критерием верности. Количественная оценка эффективности процедур кодирования определяется путем сравнения оцениваемого метода с выбранной эталонной процедурой, в качестве которой целесообразно выбирать оптимальный метод кодирования, позволяющий реализовывать потенциальные характеристики системы обработки (потенциально минимальная скорость кодирования речи таким методом, обеспечивающая восприятие речи нормального качества, составляет около 1000 бит/с). Так как оптимальный метод кодирования требует полной статистической определенности источника, не

имеющей места в реальной ситуации, и произвольной задержки сообщения, то в задачах речевого кодирования обычно за эталонную процедуру принимается стандарт кодирования речи методом импульсно-кодовой модуляции (ИКМ) со скоростью $V_{\text{ИКМ}} = 64$ кбит/с. Одним из частных показателей эффективности выступает коэффициент сжатия речевого сигнала Φ , равный отношению скоростей кодирования речи сравниваемыми алгоритмами:

$$\Phi = V_{\text{ИКМ}}/V_{\text{к}}.$$

Общепринятая система классификации методов кодирования РС подразделяет их на методы непосредственного кодирования (методы кодирования формы РС), методы вокодерного кодирования и методы гибридного кодирования. Соответствующие речепреобразующие устройства, выполняющие преобразование РС в цифровую форму на передающей стороне (кодеры) и восстановление его аналоговой формы на приемной стороне (декодеры), получили название кодеков формы сигнала (КФС), вокодеров и гибридных кодеков. Особенностью гибридных кодеков является одновременное использование принципов непосредственного и вокодерного кодирования. Следует иметь в виду, что ранее речевые кодеки подобного типа часто назывались полувокодерами.

Конечные совокупности точно сформулированных правил реализации методов кодирования РС вместе с указанием порядка их применения формируют множество алгоритмов кодирования речи.

5.4.2. Методы кодирования формы волны речевого сигнала

Методы кодирования формы волны, требующие наибольшие скорости передачи РС, до сих пор находят широкое применение, и на их основе разработаны многие международные стандарты передачи речи (рис. 5.10). Временные методы непосредственного кодирования РС осуществляют преобразование в цифровой вид формы сигнала во временной области. Импульсно-кодовая модуляция (ИКМ) основана на временной дискретизации аналогового РС в соответствии с теоремой В.А. Котельникова и дальнейшей оцифровке полученных мгновенных значений посредством квантования отсчетов и кодирования номера уровня квантования.

Кодирование РС методами дифференциальной ИКМ (ДИКМ) и дельта-модуляции (ДМ) характеризуются дополнительным включением в процесс обработки сигнала процедуры предсказания текущего отсчета РС линейной комбинацией определенного количества предыдущих отсчетов. При этом по каналу связи передается не мгновенное значение РС, а кодированная разность между его дейст-



Рис. 5.10. Классификация методов непосредственного кодирования РС

вительным и предсказанным значениями с учетом ее полярности при ДИКМ или только информация о полярности этой разности при ДМ. Многовариантность основных процедур — дискретизации, предсказания, квантования и кодирования — обуславливает разнообразие алгоритмов кодирования РС на основе временных методов. Признаками дальнейшей классификации алгоритмов выступают, как правило, режимы выполнения первых трех процедур (фиксированные или адаптивные), шкалы дискретизации, квантования и кодирования (равномерные или неравномерные), вид функции предсказания (линейная или нелинейная), способ квантования (скалярное или векторное).

Временные методы непосредственного кодирования РС имеют большое число преимуществ, в том числе относительную простоту реализации и высокое качество речи, и достаточно активно применяются в телекоммуникационных приложениях. Международным союзом электросвязи рекомендованы к применению в телефонных сетях общего пользования алгоритмы импульсно-кодовой модуляции по стандарту G.711 со скоростью кодирования 64 кбит/с и адаптивной ДИКМ (АДИКМ) по стандартам G.721, G.726, G.727 со скоростями кодирования 40, 32, 24 и 16 кбит/с. Основным недостатком цифрового сигнала, полученного методом ИКМ, является существенная статистическая избыточность, сохраняющаяся в условиях дискретизации РС, характеризующегося нестационарностью и недостаточностью априорных сведений о его статистике, с фиксированной частотой. Некоторому сокращению избыточности цифрового

сигнала при ДИКМ и ДМ, проявляющемуся в уменьшении требуемых информационных ресурсов, способствует использование декоррелирующей процедуры предсказания.

Переход из временной области описания РС в частотную позволяет по-иному использовать статистические особенности сигнала. Наиболее известными частотными методами непосредственного кодирования РС являются полосные и трансформные методы. При полосном кодировании РС подвергается спектральному анализу с помощью гребенки полосовых фильтров с амплитудно-частотными характеристиками, близкими к прямоугольным, и фазочастотными характеристиками, близкими к линейным. Разрешающая способность кратковременного спектрального анализа позволяет при этом оценить огибающую СПМ РС. Оцифрованные значения огибающей в парциальных каналах в совокупности составляют кодовое слово кодека. В простейшем случае информационные ресурсы распределяются между анализируемыми частотными полосами фиксированно с учетом сегментно-спектральной характеристики слухового восприятия, в более совершенных кодеках — с адаптацией к характеру кратковременной СПМ. В декодере полосного кодека РС воспроизводится с помощью синтезирующей гребенки полосовых фильтров на основе декодированных сигналов парциальных каналов. Следует иметь в виду, что в большинстве частотных методов непосредственного кодирования не используются фазовые соотношения частотных составляющих РС.

При трансформном кодировании информационными параметрами, передаваемыми по каналу связи, являются коэффициенты разложения сигнала по определенной системе базисных функций, в числе которых наиболее часто используются тригонометрические и комплексные экспоненциальные функции, функции Уолша и Хаара.

Известны реализации речевых кодеков, основанных на полосном и трансформном методах непосредственного кодирования, для скоростей кодирования от 8 до 32 кбит/с, однако по совокупности своих характеристик рассматриваемые методы не получили широкого применения в современных ИТКС, международные стандарты кодирования РС этими методами отсутствуют.

Как известно, в телефонии для оцифровки речи используется *импульсно-кодовая модуляция* (ИКМ, PCM — Pulse Code Modulation), рекомендация ITU-T G.711. Процесс базируется на теореме Котельникова, в соответствии с которой цифровой сигнал, полученный из аналогового сигнала выборкой, частота которой в два раза выше максимальной частоты спектра сигнала, можно обратно преобразовать в аналоговую форму. Человеческая речь воспроизводится

с приемлемым качеством в полосе частот 200...3400 Гц, чему соответствует частота выборки 8 кГц. После этого дискретизированное сообщение подвергается квантованию по q уровням, в результате чего каждому значению выборки $S(n)$ ставится в соответствие число $S_q(n)$, представленное b -разрядной комбинацией двоичного кода. Для достижения качества восприятия восстановленного речевого сигнала при равномерном квантовании необходимо, чтобы $b \geq 12$. Столь большое число уровней квантования $q = 2^{12}$ при $f_d = 8$ кГц требует, чтобы скорость I передачи символов в канале была не менее 96 кбит/с.

Однако в связи с тем, что восприятие речи человеческим ухом менее чувствительно к искажениям в области больших мгновенных значений $S(n)$, чем малых (свойство маскировки), допустимы значительно большие искажения больших мгновенных значений $S(n)$, чем малых. Поэтому требуемое число уровней квантования может быть существенно снижено в результате использования неравномерного квантования. Так, можно компрессировать либо исходный сигнал по логарифмическому закону с последующим равномерным квантованием при сравнительно малом числе уровней (например, при $b = 8$), либо предварительно преобразованный в цифровую форму сигнал при сравнительно большом исходном числе уровней квантования (при $b = 12$) с последующим преобразованием к восьмиразрядному коду (рис. 5.11).

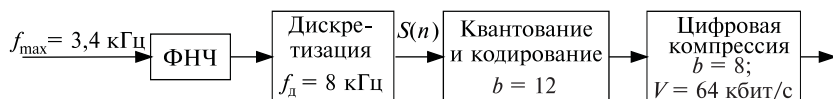


Рис. 5.11. Вариант схемы кодирования речи по Рекомендации G.711

Для обработки входных речевых сигналов с большим динамическим диапазоном используются два закона сжатия, называемые импульсно-кодовой модуляцией по закону А-ИКМ и по закону μ -ИКМ. Характеристики А-ИКМ и μ -ИКМ существенно не изменяются и остаются сравнительно постоянными в большом диапазоне уровней входного сигнала. По сравнению с простыми квантователями квантователи сжатия требуют меньше битов на входную выборку для определенного динамического диапазона сжатия. В квантователях сжатия уровни квантования находятся в области малых амплитуд, которые увеличиваются при увеличении диапазона входного сигнала. Благодаря этому при квантовании речевых сигналов, у которых максимум функции распределения вероятностей находится

в начале координат, наиболее часто встречающиеся малые амплитуды квантуются с большей точностью, чем менее вероятные большие амплитуды, что приводит к значительно лучшим, чем у простого квантователя, характеристикам.

Сжатие по A -закону определяется зависимостью:

$$y_A(\lambda) = \begin{cases} \frac{A\lambda}{1 + \lg A} & \text{при } 0 \leq \lambda \leq \frac{1}{A}; \\ \frac{1 + \lg(A\lambda)}{1 + \lg A} & \text{при } \frac{1}{A} \leq \lambda \leq 1, \end{cases} \quad (5.8)$$

где A — параметр сжатия с типовыми значениями 87,56 (Европейская ИКМ) для речевых квантователей.

Сжатие по μ -закону определяется выражением:

$$y_\mu(\lambda) = \text{sign}(\lambda) \frac{V_0 \lg(1 + \mu|\lambda|/V_0)}{\lg(1 + \mu)},$$

где V_0 задается формулой $V_0 = L\sigma_\lambda$, в которой L — нагрузочный фактор, а σ_λ — среднеквадратическое значение речевого сигнала.

Типовое значение фактора сжатия μ равно 255. Выражение (5.8) показывает, что A -закон — это комбинация логарифмической кривой, используемой для больших амплитуд, и линейного участка, используемого на малых амплитудах, μ -закон не является в точности линейным или логарифмическим ни в одном диапазоне, однако является приблизительно линейным для малых амплитуд и приблизительно логарифмическим — для больших.

Наряду с ИКМ применяются и более эффективные цифровые методы передачи речи. В частности, с целью снижения требований к пропускной способности канала можно использовать наличие корреляции между значениями отсчетов передаваемого сообщения, что используется в *дифференциальной импульсно-кодовой модуляции* (ДИКМ, DPCM — Differential Pulse Code Modulation). Отличительным признаком ДИКМ является использование в качестве единственного информационного параметра, передаваемого в цифровом виде по каналу связи, квантованной разности между текущим отсчетом дискретного РС $s(n)$ и его предсказанным значением $s_{\text{пр}}(n)$ (квантованного сигнала ошибки предсказания), что обуславливает включение в состав кодера предсказателя. Такой подход значительно сокращает динамический диапазон квантуемого сигнала. Изначально предлагалось выполнять кодирование разности между текущим и предшествующим отсчетами РС, однако введение схем предсказания позволило еще в большей степени уменьшить диапазон возможных значений сигнала, подвергаемого квантованию.

Структурная схема кодера дифференциальной импульсной модуляции с фиксированными процедурами квантования и предсказания представлена на рис. 5.12. Стандартная процедура перехода от аналогового РС $s(t)$ к дискретному (а на практике — к цифровому) сигналу $s(n)$ на рисунке не показана.

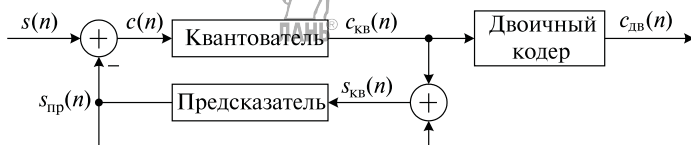


Рис. 5.12. Структурная схема кодера дифференциальной импульсно-кодовой модуляции

Предсказатель по текущему и предыдущим отсчетам квантованного дискретного речевого сигнала $s_{кв}(n)$ формирует прогнозную оценку $s_{пр}(n)$. На практике достаточно часто используются линейные предсказатели, формирующие текущий отсчет сигнала $s_{пр}(n)$ в виде линейной комбинации предыдущих отсчетов сигнала $s_{пр}(n)$; такой метод получил название метода линейного предсказания и представлен в разд. 5.3.

Сигнал ошибки предсказания определяется разностью оригинального и предсказанного сигналов:

$$c(n) = s(n) - s_{пр}(n).$$

Скалярный квантователь формирует квантованный сигнал ошибки предсказания $c_{кв}(n)$, который в двоичном виде $c_{дв}(n)$ поступает в канал связи. Сигнал ошибки квантования $e(n)$ определяется выражением $e(n) = c_{кв}(n) - c(n)$.

Таким образом, в результате совместного функционирования квантователя и предсказателя оригинальный РС $s(n)$ заменяется на его копию $s_{кв}(n)$, представляющую собой сумму предсказанного сигнала $s_{пр}(n)$ и квантованного сигнала ошибки предсказания $c_{кв}(n)$:

$$s_{кв}(n) = s_{пр}(n) + c_{кв}(n).$$

Основное преимущество метода ДИКМ над известным методом ИКМ состоит в следующем. При обеспечении работы предсказателя, близкой к оптимальной, формируемый им сигнал $s_{пр}(n)$ обладает высокой степенью похожести с оригинальным сигналом $s(n)$. В этом случае сигнал ошибки предсказания $c(n)$, квантуемый и кодируемый для передачи по каналу связи, имеет дисперсию (мощность), значительно меньшую по сравнению с мощностью исходного сигнала $s(n)$, передаваемого по каналу связи кодером ИКМ. Такое

уменьшение позволяет, не снижая точности передачи РС (точности квантования), уменьшить объем выходного алфавита квантователя (количество уровней квантования), а следовательно, понизить скорость кодирования.

Структурная схема декодера ДИКМ с фиксированной процедурой предсказания, входящего в состав приемного оборудования, представлена на рис. 5.13, при этом предсказатели кодера и декодера идентичны.

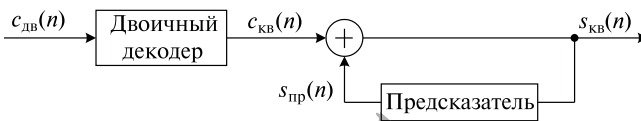


Рис. 5.13. Структурная схема декодера дифференциальной импульсно-кодовой модуляции

Отличие сигнальной дискретной копии $s_{кв}(n)$, доставляемой абоненту системы телефонии на основе ДИКМ, от оригинального дискретного сигнала $s(n)$ определяется сигналом ошибки квантования:

$$s_{кв}(n) - s(n) = s_{пр}(n) + c_{кв}(n) - s_{пр}(n) - c(n) = c_{кв}(n) - c(n) = e_{кв}(n),$$

а следовательно, зависит как от точности процедуры предсказания, определяющей мощность сигнала $c(n)$, так и от точности процедуры квантования, определяемой способом ее реализации и объемом алфавита выходных значений квантователя (допустимой разрядностью двоичного представления отсчетов сигнала $c_{кв}(n)$ или допустимой скоростью выходного цифрового потока кодера ДИКМ).

Для снижения погрешности передачи РС параметры квантователя и предсказателя должны быть согласованы со статистическими характеристиками сигнала. Поскольку РС является нестационарным процессом, то указанные процедуры должны быть адаптивными, что означает регулярную подстройку параметров квантователя и предсказателя под текущие статистические характеристики кодируемого сигнала. Адаптивность процедур квантования и (или) предсказания является классификационным признаком *адаптивной* ДИКМ (АДИКМ, ADPCM) рекомендации ITU-T G.721 и G.726.

Адаптация квантователя (как правило, в виде изменения шага квантования при сохранении количества уровней квантования) производится на основании оценки дисперсии текущего сегмента РС, при этом оценка может осуществляться на основе анализа входного или выходного сигнала квантователя. В первом случае реализуется

прямое управление квантователем (управление по входу), во втором — обратное управление (управление по выходу).

Адаптация предсказателя заключается в регулярном изменении его параметров таким образом, чтобы обеспечивался минимум дисперсии ошибки предсказания $s(n)$.

В основе дельта-модуляции (ДМ) лежит определение знака приращения сигнала за период дискретизации. ДМ можно рассматривать как частный случай ДИКМ, при котором ошибка предсказания квантуется только на два уровня, тогда при передаче в каждый момент отсчета используется всего один бит. Применение одного разряда для представления ошибки предсказания приводит к специфическим особенностям (рис. 5.14):

- при кодировании сигналов постоянного уровня аппроксимируемый сигнал «скачет» относительно кодируемого уровня (гранулярный шум);
- при кодировании быстроизменяющихся сигналов появляются ошибки, обусловленные невозможностью изменения аппроксимируемого сигнала более, чем на один шаг квантования (перегрузка по крутизне).

Для достижения на приемной стороне качества формируемого РС, сравнимого с получаемым при ИКМ G.711, необходимо значительное увеличение частоты дискретизации исходного РС. На практике классическая ДМ практически не используется, а применяются ее разновидности:

- ДМ с задержкой;
- дельта-сигма-модуляция;
- дельта-модуляция с двойным интегрированием.

При использовании адаптивной дельта-модуляции на основе анализа передаваемой информации меняется шаг квантования. Например, в ДМ с непрерывно изменяющейся крутизной (ДМНИК) информация о размере шага квантования извлекается из передаваемого цифрового потока: если подряд идут четыре нуля или единицы, шаг квантования уменьшается/увеличивается вдвое. Эта идея реализована в серийно выпускаемых микросхемах со скоростью передачи данных 12 и 16 кбит/с и характеризуется достаточно хорошей разборчивостью и узнаваемостью.

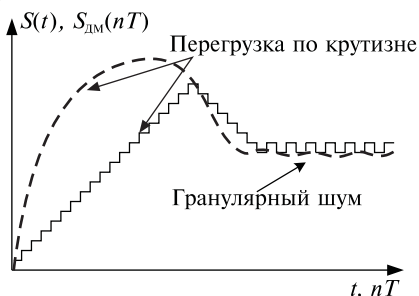


Рис. 5.14. Специфические ошибки при дельта-модуляции

5.4.3. Вокодерное кодирование

Современное состояние теории и практики речевого кодирования характеризуется доминированием речепреобразующих устройств, называемых вокодерами. Вокодерные методы речевого кодирования нацелены на достижение более значительной степени сжатия РС, чем в кодах формы сигналов. Сутью вокодерного кодирования является отказ от передачи по каналу связи информации о форме сигнала, выделение из анализируемого сегмента речи характерных признаков (параметров) или элементов, позволяющих на приемной стороне синтезировать РС. По этой причине достаточно часто кодирующее устройство вокодера называется анализатором, а декодирующее — синтезатором. Развитие вокодерных методов кодирования соответствовало процессу исследования структуры РС на акустическом, фонетическом и лингвистическом уровнях, на сегодняшний день верхней ступенью классификационной иерархии вокодерных методов является разделение их на речезлементные и параметрические методы (рис. 5.15).

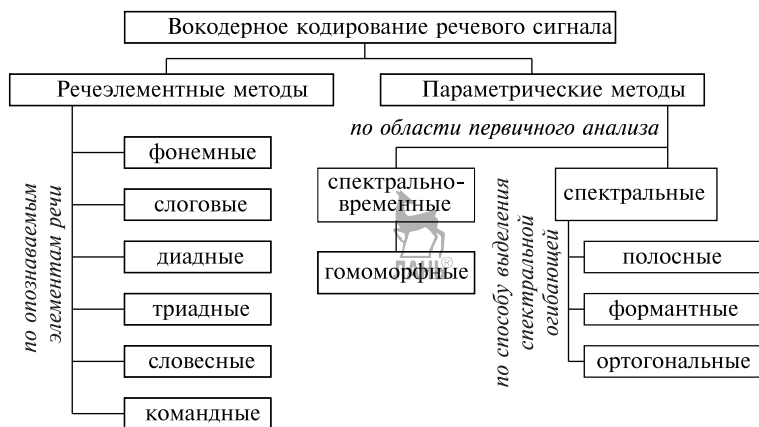


Рис. 5.15. Классификация методов вокодерного кодирования РС

Анализаторы речезлементных вокодеров распознают произнесенные элементы речи, затем производятся кодировка и передача номеров этих элементов по каналу связи. Опознаваемый единичный элемент речи (фонема, слог, группа слогов, слово, команда) служат признаком классификации речезлементных вокодеров. На приемной стороне требуемый элемент речи синтезируется по правилам речеобразования или извлекается из устройства памяти. Речезлементные вокодеры обеспечивают значительную степень сжатия РС, в некоторых случаях — близкую к потенциальной, однако не обеспечивают

требуемые абонентами натуральность и узнаваемость речи. По этой причине вокодеры данного типа не нашли широкого применения в современных телекоммуникационных системах, но активно используются в области речевого управления.

Параметрические методы вокодерного кодирования реализовывают выделение на этапе анализа медленно изменяющихся параметров речевого сигнала, обеспечивающих синтез речи в декодере. Область первичного анализа для выделения кодируемых параметров речи (сегменты непрерывного сигнала или последовательность дискретов РС) классифицируют параметрические методы на спектральные и спектрально-временные. Дальнейшее разделение спектральных методов на полосные, формантные и ортогональные основано на используемом способе описания спектральной огибающей: в виде ординат спектра на ряде частот (узких полос) речевого диапазона, совокупностью формантных параметров или суммой ортогональных функций, составляющих огибающую спектра. Очевидно, что в первом случае огибающая спектра РС аппроксимируется ступенчатой кривой, во втором — резонансными кривыми, а в третьем — совокупностью масштабированных базисных функций.

В полосных вокодерах (рис. 5.16) спектр сигнала делится на несколько, как правило 7–20 полос (каналов), при помощи полосовых фильтров. При делении на большое число каналов возрастает качество синтезированной речи наряду с увеличением скорости передачи. На выходе каждого ПФ выделяется огибающая сигнала в данной полосе частот, которая характеризует энергию РС в этой полосе. РС поступает также на устройство выделения основного тона

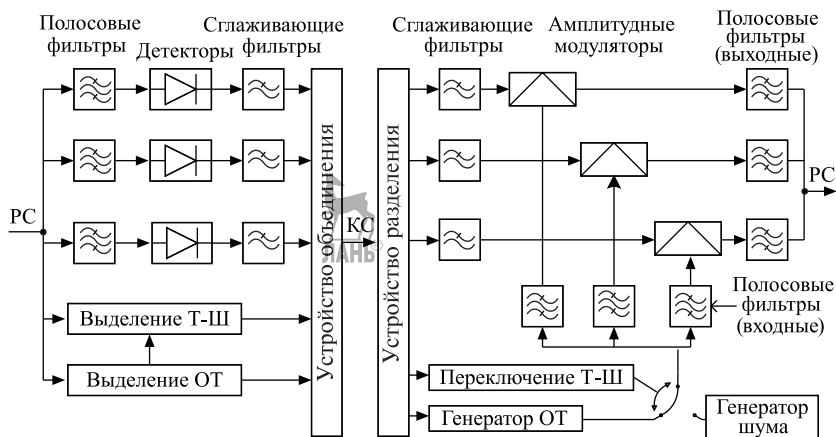


Рис. 5.16. Полосный вокодер

(ОТ) и анализатор тон-шум (Т-Ш), на выходе которых формируется сигнал, отражающий частоту основного тона и сигнал, характеризующий характер спектра (вокализованный звук — невокализованный) соответственно. Все три вида сигналов объединяются в единый информационный поток и передаются по каналу связи.

На приеме сигнал подвергается обратному преобразованию. Сигнал тон-шум управляет переключателем, который подает на гребенку синтезирующих фильтров либо импульсную последовательность (частоту которой определяет сигнал ОТ на передающем конце), либо широкополосный шум от генератора шума. Полосовые вокодеры синтезируют РС невысокого качества при относительно низкой скорости передачи.

В формантных вокодерах спектральная огибающая РС описывается комбинацией формант (резонансных частот голосового тракта). Основными параметрами формант являются центральная частота, уровень и ширина ее спектра на уровне -3 дБ. В формантном вокодере происходит выделение из РС управляющих сигналов (сигнал-параметров), которые на приеме воздействуют на резонансные контуры и воспроизводят требуемую огибающую спектра. Установлено, что при небольшом ущербе для индивидуальности звучания можно ограничиться передачей только частот и амплитуд трех формант, что сделано в большинстве формантных вокодеров.

Работа ортогональных вокодеров основана на том, что РС можно промоделировать откликом на возбуждающий сигнал линейной системы с некоторой заданной импульсной характеристикой с переменными параметрами, так что выходной сигнал равен свертке возбуждающего сигнала и импульсного отклика голосового тракта, при условии, что форма голосового тракта неизменна. В результате огибающая мгновенного спектра раскладывается в ряд по выбранной системе ортогональных базисных функций. Вычисленные коэффициенты этого разложения передаются на приемную сторону. Наибольшее распространение получили гармонические вокодеры, использующие разложение в ряд Фурье.

Современные спектрально-временные вокодеры достаточно часто в качестве объекта анализа используют РС, оцифрованный методом ИКМ, и по этой причине называются транскодеками. Основным признаком классификации спектрально-временных вокодеров выступает способ выделения параметров. Наибольшую известность среди данного класса методов приобрел метод гомоморфной фильтрации.

Параметрические методы вокодерного кодирования речи обеспечивают существенное сжатие РС, чем значительно превосходят мето-

ды непосредственного кодирования. Однако синтезированная речь на выходе параметрических вокодеров, как правило, имеет худшие, чем на выходе КФС, качество, натуральность и узнаваемость. Схемные реализации вокодерных систем значительно сложнее и дороже, многие из вокодерных методов требуют значительного времени для выполнения анализирующей процедуры. Совокупность параметров АЦП, свойственная конкретным типам параметрических вокодеров, делает невозможным использование большинства из них в современных информационно-телекоммуникационных системах.

5.5. Гибридное кодирование речевого сигнала на основе метода линейного предсказания

Доминирующее положение в классе устройств речеобразования, использующихся в современных инфокоммуникационных системах, занимают гибридные кодеки, реализующие метод линейного предсказания речи и использующие в качестве информационных параметров состояние голосового тракта (спектральную огибающую, фильтрующую функцию) и сигнала возбуждения (генераторную функцию). Данный тип кодеков получил название липредеров и сочетает в себе преимущества методов непосредственного и параметрического вокодерного кодирования.

5.5.1. Основные процедуры гибридного кодирования речевого сигнала

Принципиально важно, что в основе липредерного кодирования лежит локально-стационарная модель речеобразования, представляющая текущий сегмент РС в виде реакции линейной дискретной динамической системы (цифрового фильтра, ЦФ) на соответствующий сигнал возбуждения и предполагающая посегментовое кодирование параметров ЦФ, моделирующего голосовой тракт человека (фильтрующая функция), и сигнала возбуждения, моделирующего исходный турбулентный шум (генераторная функция). Длительность участка локальной стационарности модели речеобразования не должна превышать длительности участка квазистационарности РС. Анализатор речеобразующего устройства выделяет из короткого сегмента речевого сигнала параметры ЦФ и сигнала возбуждения, позволяющие синтезатору восстановить исходный сигнал с требуемой степенью верности, и эффективно представляет их в цифровой форме; при этом совместно решаются задачи сжатия РС и его преобразования из аналогового вида в цифровой. Переменный характер состояния системы нацелен на повышение эффективности передачи РС за счет доступной степени использования нестационарных свойств речи.

Простейшая структурная схема системы передачи РС с использованием параметрического кодирования на основе метода линейного предсказания для случая идеального канала связи с передаточной функцией $T(z) = 1$ представлена на рис. 5.17.

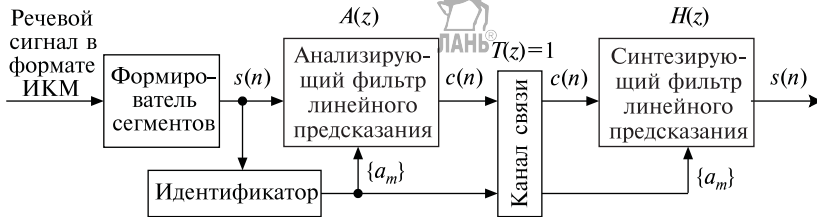


Рис. 5.17. Простейшая структурная схема системы передачи речевого сигнала на основе липредерного кодирования при идеальном канале связи

Информационными параметрами РС при его кодировании на основе метода линейного предсказания служат множество параметров линейного предсказания $\{a_m\}$ и сигнал ошибки предсказания $c(n)$, выделенные из исходного РС на этапе анализа. Первый из названных параметров присущ параметрическому кодированию, а второй содержит информацию о форме РС, что и объясняет классификацию липредерного метода кодирования как гибридного. Многовариантность определения, комбинирования и отображения указанных параметров является причиной разнообразия существующих и предлагаемых алгоритмических решений липредеров.

Практическая реализуемость рассматриваемого метода кодирования и желательность сокращения пропускной способности канала связи, необходимой для передачи оцифрованной речи, требуют квантования параметров $\{a_m\}$ и $c(n)$ перед их передачей по каналу связи. Естественно, что в этом случае даже при идеальном канале связи сегмент синтезированного РС $\hat{s}(n)$, формируемый на выходе синтезирующего фильтра, будет отличаться от исходного сегмента $s(n)$. Существенно, что в практике гибридного кодирования РС наравне со скалярным квантованием активно используется векторное квантование, особенности которого будут показаны несколько ниже. На рис. 5.18 показаны изменения структурной схемы системы передачи РС на основе липредерного кодирования (см. рис. 5.17) на основе учета процедуры квантования.

Таким образом, к основным процедурам передачи и приема РС с использованием липредерного кодирования относятся последовательно выполняемые:

- разделение дискретного РС $s(n)$ неопределенной длительности на сегменты фиксированной длины, $s(n)$, $n = 0, 1, 2, \dots, N - 1$;



Рис. 5.18. Структурная схема части системы передачи речевого сигнала на основе липредерного кодирования при идеальном канале связи с учетом процедур квантования параметров, передаваемых по каналу связи

- параметрическая идентификация фильтров ЛП на текущем сегменте РС, заключающаяся в расчете оптимальных коэффициентов линейного предсказания $\{a_m\}$;
- анализирующая цифровая фильтрация сегмента РС на основе рассчитанных коэффициентов $\{a_m\}$ с целью определения сигнала ошибки предсказания $c(n)$, $n = 0, 1, 2, \dots, N - 1$;
- квантование (скалярное или векторное) и двоичное кодирование параметров $\{a_m\}$ и $c(n)$;
- передача квантованных параметров $\{\hat{a}_m\}$ и $\hat{c}(n)$ в виде двоичного кодового слова по цифровому каналу связи;
- синтезирующая цифровая фильтрация принятого сигнала $\hat{c}(n)$, $n = 0, 1, 2, \dots, N - 1$, на основе принятых коэффициентов $\{\hat{a}_m\}$ для формирования квантованного сегмента речевого сигнала $\hat{s}(n)$, $n = 0, 1, 2, \dots, N - 1$;
- формирование синтезированного РС $\hat{s}(n)$ на основе «сборки» последовательно поступающих сегментов (на рис. 5.17 и 5.18 данная процедура не отображена).

Ниже представлены особенности выполнения первых четырех из перечисленных процедур в современных липредерах.

1. Исходный РС рассматривается как совокупность отрезков стационарных случайных последовательностей, каждый из которых порождается в системе с постоянными параметрами. Вычислительные приемы позволяют снизить нежелательное влияние переходных процессов, возникающих при переходе от одного сегмента к другому. Такой подход основан на гипотезе о локально-стационарном характере процесса речеобразования, которая позволяет оценивать изменения параметров или вероятностных характеристик нестационарного РС на основе моделей, структурно инвариантных к временному сдвигу. Длительность сегмента РС определяется длительностью участка квазистационарности РС и в большинстве составляет от 10 до 30 мс.

2. Локальная стационарность анализирующего и синтезирующего фильтров ЛП на длительности сегмента РС относит задачу оценки параметров $\{a_m\}$ по наблюдаемой реализации $s(n)$ к классу задач параметрической идентификации авторегрессионных стационарных систем во временной области. В качестве настраиваемой модели выступает рекурсивный цифровой фильтр, синтезирующий РС.

Необходимо отметить, что множество коэффициентов линейного предсказания $\{a_m\}$ не является единственным множеством, описывающим состояние фильтров ЛП на сегменте анализа. Во многих липредерных алгоритмах речепреобразования используются другие, математически эквивалентные коэффициентам ЛП, но отличающиеся устойчивостью к ошибкам квантования и канальным ошибкам, множества параметров линейного предсказания, среди которых — коэффициенты отражения (коэффициенты частичной корреляции), коэффициенты отношения площадей двух соседних секций голосового тракта, линейные спектральные частоты и т. д.

3. Следует иметь в виду, что значительная часть современных липредеров использует метод *анализа через синтез* (AbS), что устраняет из приведенного выше перечня основных процедур передачи и приема РС с использованием липредерного кодирования процедуру анализирующей цифровой фильтрации сегмента РС. В этом случае на этапе разработки липредера необходимое количество образцов (векторов) сигнала возбуждения $c(n)$ помещаются в специальную кодовую книгу, заменяющую процедуру непосредственного векторного квантования.

4. Одной из особенностей современных липредеров является активное применение процедуры векторного квантования.

Известно, что процедура скалярного квантования заключается в переходе от бесконечного числа возможных значений квантуемой случайной величины x к конечному числу разрешенных значений (центроидов) x_i , $i = 1, 2, 3, \dots, I$. В этом случае область изменения квантуемой величины разделяется на I кластеров (рис. 5.19), имеющих, в случае неравномерного квантования различные размеры, а в каждом кластере назначается центроид x_i .

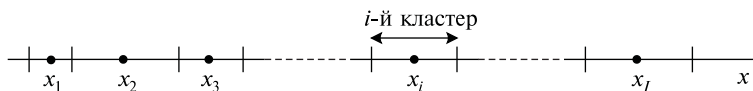


Рис. 5.19. Графическое представление процедуры скалярного квантования

При квантовании текущего значения x по каналу связи в виде двоичной комбинации передается лишь номер i кластера (центроида), соответствующего квантуемому значению x .

Аналогично в случае векторного квантования Q -мерного случайного вектора $\mathbf{x} = (x_1, x_2, x_3, \dots, x_q, \dots, x_Q)$ (в качестве таких векторов при липредерном кодировании сегмента речевого сигнала может выступать M -мерный вектор коэффициентов линейного предсказания $\mathbf{a} = (a_1, a_2, a_3, \dots, a_M)$ или N -мерный вектор сигнала возбуждения $\mathbf{c} = \{c(0), c(1), c(2), \dots, c(N-1)\}$) Q -мерное пространство с бесконечным числом возможных векторов \mathbf{x} разделяется на I кластеров (в общем случае разных форм и размеров), в каждом из которых по определенному правилу назначается вектор-центроид \mathbf{x}_i , $i = 1, 2, 3, \dots, I$.

Таким образом, вместо раздельного скалярного квантования каждой координаты x_q вектора $\mathbf{x} = (x_1, x_2, x_3, \dots, x_q, \dots, x_Q)$ применяется совместное одновременное квантование всех координат. При квантовании текущего значения случайного вектора \mathbf{x} по каналу связи в виде двоичной комбинации передается лишь номер i -го кластера, к которому принадлежит квантуемый вектор \mathbf{x} (номер центроида $\hat{\mathbf{x}}_i$). На рис. 5.20 представлена графическая иллюстрация векторного квантования для двумерного случая ($Q = 2$), жирной линией на рисунке ограничена область изменения вектора \mathbf{x} .

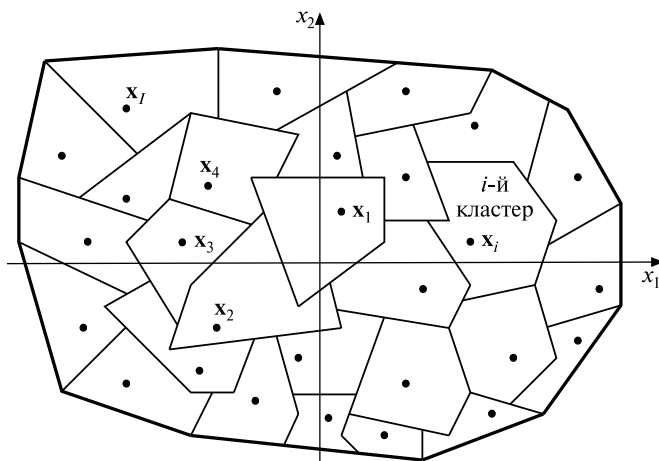


Рис. 5.20. Графическое представление процедуры векторного (двумерного) квантования

Количество кластеров I определяется на основе компромисса между допустимой скоростью кодирования (желательным коэффициентом сжатия) и требуемым качеством восстановления сигнала.

Естественно, что при обоих способах квантования приемная и передающая части аппаратуры связи используют одинаковые законы квантования (правило разделения пространства сигналов на кластеры и правило назначения центроидов в кластерах).

Известно, что в общем случае скалярное квантование способно учитывать такие характеристики квантуемых векторов (их координат), как линейные зависимости координат и их форму функции плотности распределения вероятностей. Векторное же квантование позволяет дополнительно использовать нелинейные зависимости координат и многомерность квантуемого вектора, что обеспечивает большее сжатие квантуемых сигналов (более эффективное устранение избыточности), чем при использовании скалярного квантования.

Существенно, что в настоящее время липредеры доминируют в классе устройств кодирования речи в диапазоне скоростей менее 16 кбит/с. Основные причины такой ситуации:

1. Общая востребованность систем аналого-цифрового преобразования речи, обусловленная прогрессирующей цифровизацией телекоммуникационных сетей, проявляющейся в широком практическом применении техники ЦОС.

2. Высокая степень согласованности анализа речи на основе метода ЛП с природой РС.

- 2.1. Локально-стационарный характер модели речеобразования, используемой в гибридных кодеках с ЛП, достаточно хорошо соотносится с квазистационарным характером РС на временных интервалах, соответствующих режиму установившихся звуков.

- 2.2. Эффективное разделение передаточной функции голосового тракта и возбуждающей функции при анализе речи методом ЛП создает хорошие условия для сжатия РС с целью его качественного представления при низких скоростях кодирования.

- 2.3. Важной особенностью передаточной функции голосового тракта является то, что она в основном характеризуется резонансами, которые, в свою очередь, достаточно хорошо представляются с помощью полюсов. Следовательно, применение метода ЛП, использующего полюсную модель ЛДДС, позволяет сохранить многие важнейшие характеристики ПФ речевого аппарата человека. При этом потери, получаемые при переходе от полюсно-нулевой модели к полюсной, вполне компенсируются существенным упрощением алгоритмов обработки РС.

3. Высокая степень соответствия липредерных алгоритмов речепреобразования современному уровню развития теории, технологии и схмотехники ЦОС. Наличие эффективных расчетных алгоритмов реализации процедур анализа речи на основе метода ЛП обеспечивает достижение высокой точности анализа в режиме реального времени при соблюдении ограничений на вычислительную сложность обусловленных современным уровнем развития цифровых процессоров обработки сигналов.

4. Значительное превосходство аналого-цифрового преобразования речи на основе метода ЛП над другими методами в диапазоне скоростей кодирования менее 16 кбит/с. Высокая степень востребованности во многих телекоммуникационных приложениях указанного скоростного диапазона сопровождается его качественной недостижимостью для речепреобразующих устройств, основанных на методах непосредственного кодирования РС, речееlementных и параметрических вокодеров. Эффективное использование корреляционных зависимостей, характерных для РС, подвергнутого равномерной дискретизации во временной области, обеспечивает липредерам преимущества перед другими типами кодеков.

К настоящему времени в различных инфокоммуникационных приложениях используется большое число различных липредерных алгоритмов речевого кодирования, стандартизированных международными организациями или государственными ведомствами. Классификация этих алгоритмов может быть основана на следующих классификационных признаках:

- 1) порядок и размерность ЛП;
- 2) используемое множество параметров ЛП;
- 3) способ представления сигнала возбуждения;
- 4) используемая мера различия между оригинальным и синтезированным РС;
- 5) длительность временных интервалов перерасчета сигнала возбуждения и параметров синтезирующего фильтра;
- 6) способы квантования кодируемых параметров;
- 7) степень адаптации используемых процедур к характеристикам обрабатываемых речевых сегментов;
- 8) степень и механизм использования особенностей слухового восприятия человека.

В табл. 5.2 представлены наиболее популярные стандарты липредерного кодирования речи, при этом классификация типов липредеров выполнена в основном по области применения алгоритма и способу представления сигнала возбуждения. Липредеры, представленные в таблице, используют различные модификации метода ЛП речи:

- 1) с возбуждением от импульсов основного тона, часто называемый просто LPC (Linear Predictive Coding);
- 2) с многоимпульсным возбуждением — MPELP (Multi Pulse Excited Linear Prediction);
- 3) с возбуждением от остатка предсказания — RELP (Residual Excited Linear Prediction);



Таблица 5.2

Использование липредеров в инфокоммуникационных приложениях

Область применения	Стандарт				Скорость, кбит/с
	Наименование	Год	Тип липредера	Организация	
ТфОП	G.728	1992	LD-CELP	ITU-T	16
	G.729	1995	CS-ACELP		8
	G.722.2	2003	ACELP		переменная
Сотовые системы радиосвязи	GSM full-rate	1989	RPE-LTP-LPC	ETSI	13
	GSM half-rate	1994	ACELP		6,5
	GSM-EFR	1995	ACELP		13
	IS-54	1991	VSELP	TIA	7,95
	IS-96	1996	CELP		1,2; 2,4, 4,8; 9,6
	JDC full-rate	1991	VSELP	Министерство почты и связи Японии	6,7
	JDC half-rate	1993	PSI-CELP		3,45
Интернет-телефония	Поколение 3G (спецификации TS26.190 и TS26.194)	2001	ACELP	Разработчики систем мобильной связи совместно с ETSI	23,85; 23,05; 19,85; 18,25; 15,85; 14,25; 12,65; 8,85; 6,6
	G.723.1	1995	MP-MLQ/ACELP	ITU-T	5,27/6,3
	TETRA	1994	ACELP	ETSI	4,56
Спутниковые системы	Iridium		VSELP		2,4
	Global star		CELP		1,2; 2,4...9,6
Системы специального назначения	Inmarsat-aero	1989	MP-MLQ/ACELP		9,6
	FS-1015	1984	LPC	Министерство обороны США	2,4
	FS-1016	1990	CELP		4,6
	Аппаратура автоматического шифрования речи		CELP, ACELP		2,4; 4,8; 9,6

ITU-T — Сектор стандартизации Международного союза электросвязи; TIA — Ассоциация телекоммуникационной промышленности США; ETSI — Европейский институт стандартов в области телекоммуникаций;

4) с возбуждением от суммы векторов — VSELP (Vector Sum Excited Linear Prediction);

5) с возбуждением от кода — CELP (Code Excited Linear Prediction);

6) с регулярным импульсным возбуждением и долговременным линейным предсказанием — RPE-LTP (Regular Pulse Excited Long Term Prediction);

7) с возбуждением от алгебраического кода — ACELP (Algebraic Code Excited Linear Prediction);

8) с возбуждением от алгебраического кода и сопряженной структурой квантователя коэффициентов усиления сигналов возбуждения — CS-ACELP (Conjugate-Structure and Algebraic Code Excited Linear Prediction);

9) с возбуждением от кода и малой алгоритмической задержкой — LD-CELP (Low Delay Code Excited Linear Prediction);

10) с возбуждением от кода, отображающего сигнал возбуждения, синхронный с основным тоном — PSI-CELP (Pitch Synchronous Innovation — Code Excited Linear Prediction).

В настоящее время совершенствование липредерных кодеков речи достигается путем внедрения в алгоритмы кодирования все более сложных процедур анализа и синтеза, повышающих степень уменьшения статистической и психофизической избыточностей.

Анализ реализуемых и предлагаемых к реализации липредеров позволяет выделить доминирующие на сегодняшний день тенденции в совершенствовании метода аналого-цифрового преобразования речи на основе ЛПП:

1. Преимущественное использование для описания текущего состояния ЛДДС M -мерного вектора линейных спектральных пар (ЛСП), являющегося математическим эквивалентом множества коэффициентов ЛПП $\{a_m\}$, но обладающего лучшей устойчивостью к канальным ошибкам.

2. Использование различных вариантов липредеров с возбуждением от кода (CELP-кодеков). Особенностью CELP-кодеков является представление сигнала ошибки предсказания (сигнала возбуждения) $s(n)$ конечным множеством векторов возбуждения, хранящихся в кодовых книгах кодера и декодера.

3. Представление сигнала возбуждения многокомпонентной моделью. Наибольшую популярность приобретает двухкомпонентная модель, включающая в себя стохастический вектор возбуждения, кодируемый с помощью фиксированной кодовой книги, и вектор возбуждения, отображающий результат долговременного линейного предсказания, основанного на анализе динамики основного тона РС.

4. Определение сигнала возбуждения процедурой анализа через синтез (Analysis-by-Synthesis — AbS) с использованием рассчитанного на текущем сегменте состояния синтезирующего фильтра ЛПП. Анализ через синтез позволяет сформировать итоговый вектор возбуждения не только посредством минимизации ошибки квантования, но и с учетом критериев слухового восприятия речи человеком. Реализация процедуры AbS требует включения в структуру

Таблица 5.3

Основные характеристики липредерных алгоритмов кодирования

Характеристика	Алгоритм G.728	Алгоритм G.729	Алгоритм GSM
Объект кодирования	Речевой сигнал в полосе частот 0,3...3,4 кГц, дискретизированный с частотой 8 кГц и представленный в формате 8-битной линейной ИКМ		
Используемая модель аналогового речевого сигнала	Квазистационарный случайный процесс с интервалом стационарности, равным		
Длительность сегмента речевого сигнала	2,5 мс	10 мс	20 мс
Скорость кодирования, кбит/с	2,5 мс (20 отсчетов PC)	10 мс (80 отсчетов PC)	20 мс (160 отсчетов PC)
Способ определения оптимального сигнала (вектора) возбуждения	16	8	13 (5,6)
Максимальный интервал постоянства всех информационных параметров	Анализ через синтез с использованием кодовых книг сигнала возбуждения 0,625 мс, определяется частотой выбора вектора возбуждения	Непосредственная цифровая фильтрация	—
Алгоритмическая задержка, мс	0,625	5 мс, определяется частотой выбора вектора возбуждения	5 мс, определяется частотой выбора вектора возбуждения
Качество синтезированного РС по шкале MOS, балл	3,8	15	20
Вычислительная сложность алгоритма, MIPS	29	3,9	3,7
	20	20	

передающей части кодека цифрового фильтра, аналогичного синтезирующему фильтру декодера.

5. Широкое применение разновидностей векторного квантования для отображения кодируемых параметров речи при использовании внутрисегментовых и межсегментовых зависимостей последних с целью максимально возможного исключения избыточности.

6. Мультипликативное кодирование параметров, составляющих признаковое пространство кодеков, что обусловлено достаточной степенью статистической независимости ПФ голосового тракта (состояния синтезирующего фильтра) и генераторной функции (сигнала возбуждения) используемой модели речеобразования.

7. Использование специальных процедур обработки, позволяющих учитывать особенности восприятия РС органами слуха.

Ниже будут рассмотрены наиболее популярные алгоритмы липредерного кодирования РС, используемые в телефонных сетях об-

щего пользования (стандарт G.728 на скорость 16 кбит/с и стандарт G.729 на скорость 8 кбит/с), в сетях сотовой связи стандарта GSM (табл. 5.3) и наземной транкинговой радиосвязи TETRA.

5.5.2. Кодирование речевого сигнала в системе связи стандарта GSM

Система сотовой связи GSM (Global System for Mobile communication) является в настоящее время наиболее распространенной системой цифровой связи с подвижными объектами. В полноскоростном режиме кодирования PC (13 кбит/с) в системе связи GSM используется липредер с регулярным импульсным возбуждением и долговременным линейным предсказанием — RPE-LTP (Regular Pulse Excitation-Long Term Prediction).

Существенно, что в системе сотовой связи GSM реализован режим прерывистой передачи речи (DTX — Discontinuous Transmission), структурно показанный на рис. 5.21. Режим обеспечивает значительное снижение речевого трафика системы связи, используя тот факт, что коэффициент активности каждого абонента в телефонном диалоге не превышает 0,5.



Рис. 5.21. Схема взаимодействия устройств, обеспечивающих режим прерывистой передачи речевого сигнала в системе сотовой связи GSM

Важнейшим устройством кодера, обеспечивающим режим DTX, является детектор активности речи VAD (Voice Activity Detector), классифицирующий текущий сегмент РС как активную речь или паузу и управляющий отключением передатчика в паузах речи. Исследованиями подтверждено, что полное отсутствие фонового шума (в случае отключения передатчика в паузах речевого сигнала) уменьшает комфортность абонента и снижает разборчивость речи.

Для устранения указанных недостатков в декодер стандарта GSM включен формирователь комфортного шума, синтезирующий шумовой участок РС на основании информации о среднем уровне фонового шума, периодически передаваемой кодером в паузе речи (один раз на сегменте паузы длительностью 480 мс). Очевидно, что включение VAD в тракт кодирования РС несколько ухудшает качество синтезированной речи, что обусловлено наличием вероятности ошибочной классификации активной речи как паузы.

Структурная схема кодера РС стандарта GSM представлена на рис. 5.22.

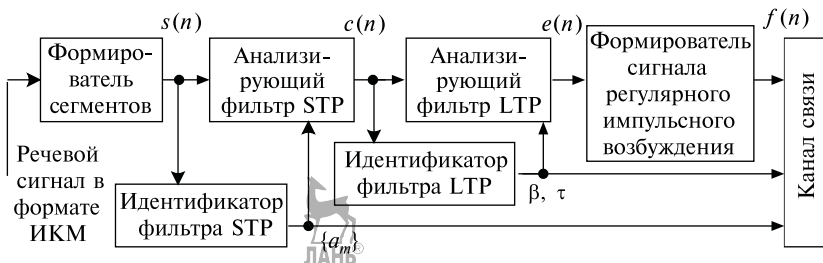


Рис. 5.22. Структурная схема кодера речевого сигнала стандарта GSM

Основные этапы кодирования РС в стандарте GSM:

1) разделение дискретного РС неопределенной длительности на сегменты фиксированной длины $s(n)$, $n = 0, 1, 2, \dots, 159$;

2) параметрическая идентификация фильтра кратковременного ЛП (STP — Short Term Prediction) 8-го порядка на текущем сегменте РС, заключающаяся в расчете оптимальных коэффициентов ЛП a_m , $m = 1, 2, \dots, 8$;

3) цифровая фильтрация сегмента РС анализирующим фильтром STP 8-го порядка на основе рассчитанных коэффициентов $\{a_m\}$ с целью определения сигнала ошибки кратковременного линейного предсказания $c(n)$, $n = 0, 1, 2, \dots, 159$;

4) параметрическая идентификация фильтра долговременного ЛП (LTP — Long Term Prediction), выполняющаяся на четверти длительности текущего сегмента РС и заключающаяся в расчете опти-

мальных задержки τT (порядка фильтра) и коэффициента долговременного ЛП β ;

5) цифровая фильтрация сигнала $c(n)$ анализирующим фильтром ЛТР с учетом параметров β и τ с целью последовательного определения (четырежды на длительности сегмента) сигнала ошибки долговременного ЛП $e(n)$, $n = 0, 1, 2, \dots, 39$;

6) формирование из сигнала $e(n)$ сигнала регулярного импульсного возбуждения $f(n)$, $n = 0, 1, 2, \dots, 39$, заключающееся в последовательном выполнении следующих процедур:

- обработки сигнала $e(n)$ цифровым фильтром нижних частот с целью получения сглаженного сигнала $e_c(n)$, $n = 0, 1, 2, \dots, 39$;
- формирования из сигнала $e_c(n)$ трех последовательностей $f_j(n)$, $j = 1, 2, 3$, $n = 0, 1, 2, \dots, 39$, ненулевые отсчеты которых определяются следующими отсчетами сигнала $e_c(n)$:

$$f_1(n) = e_c(n), \quad n = 0, 3, 6, \dots, 36;$$

$$f_2(n) = e_c(n), \quad n = 1, 4, 7, \dots, 37;$$

$$f_3(n) = e_c(n), \quad n = 2, 5, 8, \dots, 38;$$

- выбора в качестве оптимального сигнала регулярного импульсного возбуждения $f(n)$ последовательности $f_j(n)$, имеющей большую энергию;

7) квантование и двоичное кодирование параметров $\{a_m\}$, β , τ и $f(n)$; на рис. 5.21 данные процедуры не отображены (существенно, что коэффициенты кратковременного ЛП предварительно пересчитываются в математический эквивалент, называемый логарифмическим отношением площадей соседних секций голосового тракта);

8) передача квантованных параметров $\{\hat{a}_m\}$, $\hat{\beta}$, $\hat{\tau}$ и $\hat{f}(n)$ в виде двоичного кодового слова по цифровому каналу связи.

Декодер речевого сигнала стандарта GSM на основании информации, полученной из канала связи, последовательно выполняет:

- цифровую фильтрацию принятого сигнала $\hat{f}(n)$ синтезирующим фильтром долговременного ЛП с параметрами $\hat{\beta}$ и $\hat{\tau}$ с целью формирования сигнала возбуждения $\hat{c}(n)$, $n = 0, 1, 2, \dots, 159$;
- цифровую фильтрацию сигнала $\hat{c}(n)$ синтезирующим фильтром кратковременного ЛП 8-го порядка на основе принятых коэффициентов $\{\hat{a}_m\}$ с целью формирования квантованного сегмента речевого сигнала $\hat{z}(n)$, $n = 0, 1, 2, \dots, 159$;
- формирование синтезированного РС $\hat{s}(n)$ на основе «сборки» последовательно поступающих сегментов.

Сегмент РС, содержащий 160 отсчетов, отображается в полномкоротном режиме стандарта GSM 260-битовым кодовым словом.

Таблица 5.4

Распределение битов кодового слова кодека речевого сигнала стандарта GSM

Кодируемые параметры	Кол-во битов, формируемое на длительности	
	подсегмента РС (5 мс)	сегмента РС (20 мс)
Параметры фильтра STP (коэффициенты a_m в виде логарифмических отношений площадей)	—	36
Параметры фильтра LTP:		
задержка τ	7	28
коэффициент долговременного линейного предсказания β	2	8
Параметры сигнала регулярного импульсного возбуждения $f(n)$:		
номер j последовательности	2	8
максимальная амплитуда отсчета в последовательности	6	24
нормированные амплитуды отсчетов	39	156
Общее количество битов	—	260

Распределение общего количества битов по кодируемым параметрам (для сегментов активной речи) показано в табл. 5.4.

5.5.3. Метод анализа через синтез

По отношению к гибриднему кодированию на основе линейного предсказания метод анализа через синтез заключается в том, что на этапе анализа текущего сегмента РС в кодере квантованный сигнал ошибки предсказания определяется не путем анализирующей цифровой фильтрации, а на основе сравнения текущего сегмента РС с множеством его синтезированных копий. Процедура вычисления коэффициентов линейного предсказания при этом остается прежней. Структурная схема части системы передачи речевого сигнала на основе липредерного кодирования с использованием метода анализа через синтез представлена на рис. 5.23.

Сравнение рис. 5.18 и 5.23 демонстрирует отличия способа определения квантованного сигнала (вектора) возбуждения $\hat{c}(n)$ при использовании анализа через синтез от рассмотренного в п. 5.5.1:

- 1) отказ от анализирующего фильтра линейного предсказания и непосредственного векторного квантования сигнала возбуждения;
- 2) получение на этапе разработки кодека состоятельного множества вариантов сигнала ошибки линейного предсказания $\{c(n)\}$;
- 3) представление векторного квантователя в виде таблицы центроидов $\{\hat{c}_i(n)\}$, $i = 1, 2, \dots, I$, полученных на основе множества $\{c(n)\}$ (формирование «кодовой книги векторов возбуждения»);
- 4) формирование синтезирующим фильтром линейного предсказания множества сегментов $\{\hat{s}_i(n)\}$ синтезированного РС, соответствующих множеству векторов возбуждения $\{\hat{c}_i(n)\}$;

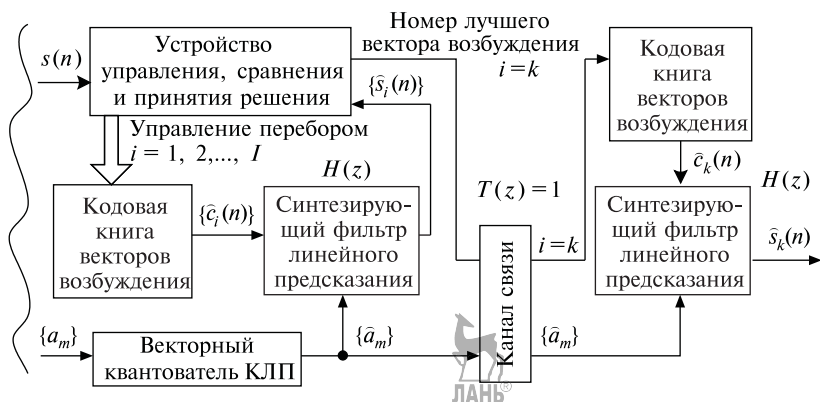


Рис. 5.23. Структурная схема части системы передачи речевого сигнала на основе липредерного кодирования с использованием метода анализа через синтез при идеальном канале связи

5) выбор устройством управления, сравнения и принятия решения сегмента $\hat{s}_i(n)$, $i = k$ синтезированного РС, максимально близкого к текущему сегменту $s(n)$ по используемому критерию;

6) передача по каналу связи номера k вектора возбуждения, обеспечивающего синтез сегмента $\hat{s}_k(n)$ речевого сигнала.

Системы гибридного кодирования речевого сигнала на основе линейного предсказания (липредерные системы), использующие метод анализа через синтез, получили название «липредеры с возбуждением от кода» — CELP-кодеки (Code Excited Linear Prediction), их разновидности представлены в табл. 5.2. Ниже представлены особенности некоторых стандартов (алгоритмов) липредерного кодирования, использующих метод анализа через синтез.

5.5.4. Кодирование речевых сигналов в системе связи стандарта TETRA

Стандарт цифровой транкинговой радиосвязи TETRA (TErrestrial Trunked RAdio) разработан европейским институтом стандартизации в области телекоммуникаций (ETSI) для систем профессиональной подвижной связи.

Стандартом TETRA предусматривается использование липредера с возбуждением от алгебраического кода (ACELP). Особенности кодека речевого сигнала стандарта TETRA являются:

- 1) использование метода анализа через синтез (см. рис. 5.23);
- 2) длительность сегмента речевого сигнала 30 мс (240 отсчетов при частоте дискретизации 8 кГц);
- 3) разделение исходного сегмента на субсегменты по 7,5 мс (60 отсчетов);

- 4) порядок фильтра синтеза $M = 10$;
- 5) применение линейных спектральных пар (линейных спектральных частот);
- 6) формирование сигнала возбуждения СФЛП в виде суммы масштабированных векторов возбуждения, содержащихся в адаптивной и алгебраической кодовых книгах.

Упрощенная схема кодера речевого сигнала стандарта TETRA приведена на рис. 5.24.

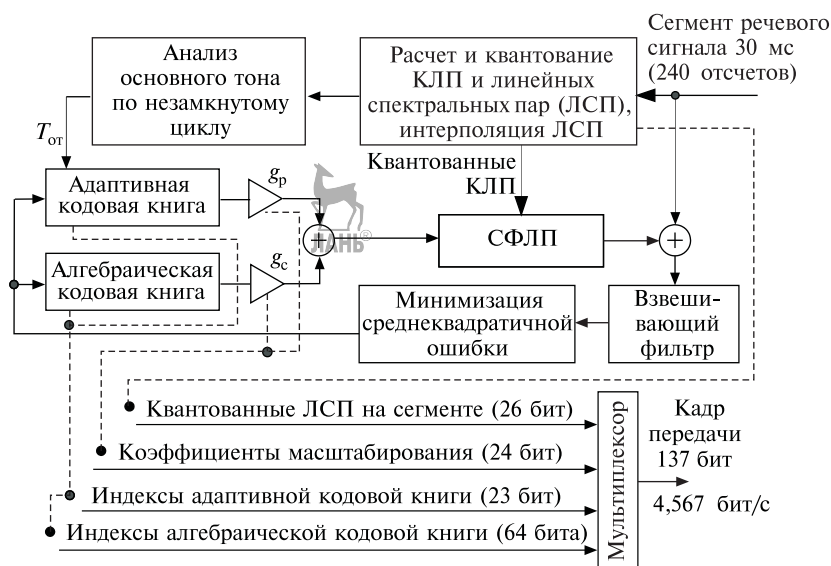


Рис. 5.24. Кодер речевого сигнала системы связи стандарта TETRA

Входной сегмент речевого сигнала длительностью 30 мс (240 отсчетов) поступает на блок вычисления 10 коэффициентов линейного предсказания (КЛП), которые преобразуются в линейные спектральные пары (ЛСП). Для повышения качества квантования десять ЛСП разделяются на три подвектора размерностью 3, 3 и 4, квантуются и интерполируются. Квантованные ЛСП описываются 26 битами в кадре передачи. Разделение вектора ЛСП на три также объясняется необходимостью повышения помехоустойчивости при передаче по каналу связи. Для функционирования фильтра синтеза используются вновь рассчитанные уже из квантованных ЛСП квантованные КЛП на длительности сегмента 30 мс.

Определение сигнала возбуждения фильтра синтеза производится на каждом субсегменте длительностью 7,5 мс (60 отсчетов).

Вычисление частоты основного тона происходит в два этапа: анализ по незамкнутому циклу и последующий поиск по замкнутому циклу при помощи адаптивной кодовой книги. На открытом цикле частота основного тона вычисляется один раз на сегменте 30 мс автокорреляционным методом. После нахождения основного тона по незамкнутому циклу выполняется анализ по замкнутому циклу для четырех субсегментов в адаптивной кодовой книге вблизи значений частоты ОТ, определенной ранее. Индексы кодовой книги для частоты ОТ четырех субсегментов кодируются; для первого выделяется 8 бит, остальные три — по 5 бит.

Как отмечалось ранее, в кодере TETRA применяется алгебраическая кодовая книга, формирующая векторы возбуждения таким образом, чтобы их энергия была сконцентрирована в наиболее важных частотных полосах. Вектор алгебраической кодовой книги содержит четыре ненулевых импульса с амплитудами $+1,4142$, -1 , $+1$ и -1 . Позиция первого импульса кодируется пятью битами, остальных — тремя битами и обновляются для каждого субсегмента. Дополнительно в кадр передачи от алгебраической кодовой книги вкладывается информация о смещении импульса и общем знаке импульса, кодируемых 1 битом каждый.

Коэффициенты усиления для векторов кодовых книг кодируются шестью битами.

Сформированный на основе двух кодовых книг сигнал возбуждения подается на СФЛП для формирования речевого сегмента, который сравнивается с исходным. Разностный сигнал поступает на взвешивающий фильтр, корректирующий ошибку предсказаний в области формант спектра речи. Оптимальные параметры формирования сигнала возбуждения СФЛП определяются по критерию минимума среднеквадратической ошибки взвешенного разностного сигнала.

Для каждого сегмента РС длительностью 30 мс в кодере ACELP стандарта TETRA формируется кадр передачи размерностью 137 бит, что эквивалентно скорости кодирования 4,567 кбит/с. Распределение информационных ресурсов по кодируемым параметрам РС представлено в табл. 5.5.

Задача декодера (рис. 5.25) — декодирование принятых параметров (коэффициентов линейного предсказания, вычисленных из принятого квантованного вектора линейных спектральных пар; векторов индексов адаптивной и алгебраической кодовых книг, коэффициентов усиления) и выполнения синтеза сегмента речевого сигнала.

Таблица 5.5

Распределение битов в кадре передачи, соответствующем сегменту анализируемого речевого сигнала в стандарте TETRA

Параметр	Подсегмент				Всего в сегменте
	1	2	3	4	
Коэффициенты линейного предсказания					26
Индекс адаптивной кодовой книги	8	5	5	5	23
Индекс алгебраической кодовой книги	16	16	16	16	64
Коэффициенты масштабирования	6	6	6	6	24
Всего					137

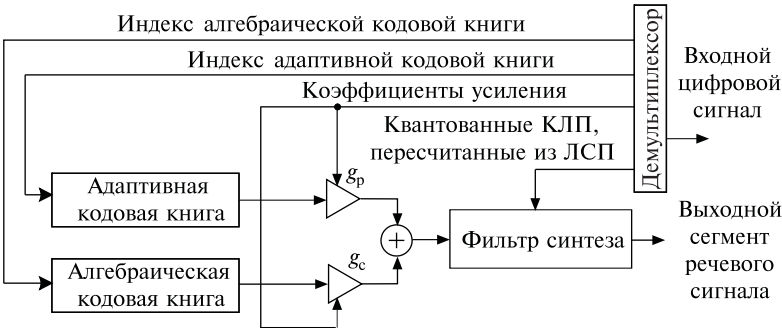


Рис. 5.25. Декодер речевого сигнала стандарта TETRA

5.5.5. Стандарты G.728 и G.729

Стандарт Международного союза электросвязи G.728 определяет алгоритм кодирования речи методом ЛП с возбуждением от кода и малой алгоритмической задержкой (LD-CELP). Структурная схема кодера, реализующего данный алгоритм, представлена на рис. 5.26.

Исходный РС в формате ИКМ делится на сегменты длительностью 2,5 мс (20 отсчетов, 160 бит), затем выполняется их последовательное независимое кодирование.

Для текущего сегмента РС первоначально определяется состояние синтезирующего фильтра (расчет его коэффициентов), затем на длительности каждой четверти сегмента (0,625 мс; 5 отсчетов РС) реализуется процедура анализа через синтез:

- формирование вариантов векторов синтезированной речи \hat{S} на основе полного перебора векторов сигнала возбуждения C , содержащихся в кодовой книге кодера;
- расчет векторов ошибки квантования $e = S - \hat{S}$ и их взвешивание фильтром восприятия, учитывающим особенности слухового аппарата человека;

Кадр исходного речевого сигнала S

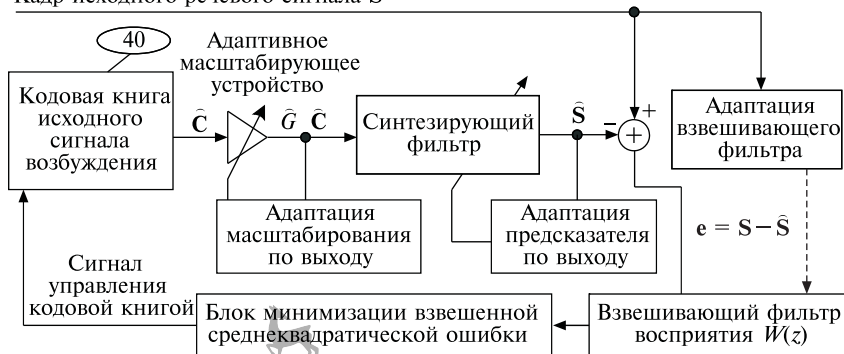


Рис. 5.26. Структурная схема кодера речевого сигнала, реализующего алгоритм G.728

- определение вектора \mathbf{C} , обеспечивающего максимальную похожесть векторов оригинального и синтезированного РС по критерию минимума взвешенной среднеквадратической ошибки.

В качестве информационного параметра, передаваемого по каналу связи от кодера к декодеру, в стандарте G.728 используется только номер лучшего вектора \mathbf{C} . Так как кодовая книга исходного сигнала возбуждения содержит 1024 варианта 5-мерных векторов \mathbf{C} сигнала возбуждения, то на каждой четверти сегмента РС формируется 10-битовое кодовое слово ($\log_2 1024 = 10$). Общее кодовое слово, отображающее текущий сегмент РС, содержит 40 бит. Этим обуславливается скорость кодирования РС в стандарте G.728:

$$V_k = 40 \frac{1}{2,5 \cdot 10^{-3}} = 16 \text{ кбит/с.}$$

Синтезирующий фильтр кодера имеет 50-й порядок, его состояние фиксируется на длительности сегмента РС, обновление происходит на третьей четверти каждого сегмента. Оптимальные коэффициенты синтезирующего фильтра определяются методом анализа на основе ЛПИ квантованного РС использованием следующих процедур:

- расчет коэффициентов автокорреляции методом взвешивания гибридной оконной функцией;
- расчет коэффициентов ЛПИ методом рекурсии Левинсона–Дарбина;
- коррекция коэффициентов ЛПИ для расширения полосы частот пиков полученного спектра с целью повышения устойчивости к канальным ошибкам.

На вход синтезирующего фильтра поступает масштабированный квантованный сигнал возбуждения $\hat{G}\hat{\mathbf{C}}$. Особенностью построения

кодовой книги исходного сигнала возбуждения \mathbf{C} является ее двухуровневая структура:

- 7-битовая книга первого уровня содержит 128 независимых пятимерных кодовых векторов;
- 3-битовая книга второго уровня содержит 8 скалярных значений, симметричных по отношению к нулю.

Результирующие векторы \mathbf{C} сигнала возбуждения представляются произведением векторов книги первого уровня на скаляры книги второго уровня.

Скалярный коэффициент масштабирования \hat{G} исходного вектора сигнала возбуждения определяется каждые 0,625 мс (на каждой четверти сегмента) с использованием процедуры адаптивного ЛП 10-го порядка с управлением по выходу.

Взвешивающий фильтр восприятия представляет собой цифровой фильтр 10-го порядка, его состояние фиксируется на длительности сегмента РС и обновляется на третьей четверти каждого сегмента. Оптимальные коэффициенты взвешивающего фильтра восприятия определяются методом анализа на основе ЛП неквантованного РС использованием следующих процедур:

- расчет коэффициентов автокорреляции методом взвешивания гибридной оконной функцией;
- расчет коэффициентов ЛП методом рекурсии Левинсона–Дарбина;
- расчет коэффициентов взвешивающего фильтра посредством пересчета коэффициентов ЛП.

К основным особенностям алгоритма кодирования РС по стандарту G.728 следует отнести малую алгоритмическую задержку РС, обуславливаемую в основном малой длительностью сегмента анализа, и использование в качестве информационных параметров, передаваемых по каналу связи, лишь номеров векторов сигнала возбуждения. Коэффициент масштабирования \hat{G} , коэффициенты синтезирующего фильтра и коэффициенты взвешивающего фильтра восприятия автономно рассчитываются в кодере и декодере липредера, а по каналу связи не передаются. Начальные состояния указанных фильтров и используемые гибридные оконные функции однозначно определены.

Характеристика информационных параметров, используемых в алгоритме G.728, представлена в табл. 5.6.

Стандарт Международного союза электросвязи G.729 определяет алгоритм кодирования речи методом ЛП с возбуждением от алгебраического кода и сопряженной структурой квантователя коэффициентов усиления сигналов возбуждения (CS-ACELP). Струк-

Таблица 5.6

Характеристика информационных параметров, используемых в алгоритме G.728

Параметр	Исходный вектор сигнала возбуждения \hat{C}	Коэффициент масштабирования исходного вектора сигнала возбуждения \hat{G}	Коэффициенты предсказания синтезирующего фильтра	Коэффициенты взвешивающего фильтра восприятия
Информационное название, способ представления	Описание сигнала поступающего на синтезирующий фильтр (генераторная функция) на длительности четверти сегмента (0,625 мс)		Фиксация состояния синтезирующего фильтра на длительности сегмента речевого сигнала (фильтрующая функция)	Фиксация состояния взвешивающего фильтра восприятия кодера на длительности сегмента речевого сигнала
	Пятимерный вектор	Скаляр	50-мерный вектор	10-мерный вектор
Способ выделения (определения), используемые процедуры	Процедура анализа через синтез, полный перебор среди 1024 векторов-кандидатов. Выбирается лучший вектор по минимуму взвешенной среднеквадратической ошибки	Адаптивное линейное предсказание 10-го порядка коэффициента масштабирования вектора возбуждения: ● расчет коэффициентов автокорреляции методом взвешивания гибридной оконной функции; ● расчет коэффициентов линейного предсказания методом рекурсии Левинсона-Дарбина; ● расширение полюсов частот полученного спектра; ● расчет предсказываемого значения G	Начальные состояния фильтров Анализ методом адаптивного линейного предсказания 50-го порядка квантованного речевого сигнала: ● расчет коэффициентов автокорреляции методом взвешивания гибридной оконной функции; ● расчет коэффициентов линейного предсказания методом рекурсии Левинсона-Дарбина; ● коррекция коэффициентов линейного предсказания для расширения полюсов частот пиков полученного спектра с целью повышения устойчивости к канальным ошибкам	Анализ методом адаптивного линейного предсказания 10-го порядка неквантованного речевого сигнала: ● расчет коэффициентов автокорреляции методом взвешивания гибридной оконной функции; ● расчет коэффициентов линейного предсказания методом рекурсии Левинсона-Дарбина; ● расчет коэффициентов взвешивающего фильтра посредством пересчета коэффициентов линейного предсказания

Окончание табл. 5.6				
Параметр	Исходный вектор сигнала возбуждения \vec{S}	Коэффициент масштабирования исходного вектора сигнала возбуждения \vec{G}	Коэффициенты предсказания синтезирующего фильтра	Коэффициенты взвешивающего фильтра восприятия
Способ квантования, используемые процедуры	Векторное квантование. Кодовая книга разбита на два уровня: 7-битовая книга (128 независимых пятимерных кодовых векторов); 3-битовая книга (8 скалярных значений, симметричных по отношению к нулю). Исходный вектор сигнала возбуждения представляется произведением вектора первой книги на скаляр второй	–	–	–
Передача по каналу связи	10-битовое кодовое слово на каждом подсегменте (0,625 мс), определяющее порядковый номер лучшего исходного вектора сигнала возбуждения	Не передаются, автономно рассчитываются в кодере и декодере липсидера		

турная схема кодера, реализующего данный алгоритм, представлена на рис. 5.27. На схеме показано количество бит, генерируемых отдельными структурными элементами кодера для формирования итогового кодового слова, отображающего анализируемый (текущий) сегмент РС.

Исходный РС в формате ИКМ делится на сегменты длительностью 10 мс (80 отсчетов, 640 бит), затем осуществляется их последовательное независимое кодирование. Для текущего сегмента РС первоначально определяется состояние синтезирующего фильтра 10-го порядка (расчет его коэффициентов) последовательным выполнением следующих процедур:

- определение коэффициентов ЛП корреляционным методом на основе анализа сегмента оригинального РС;
- пересчет коэффициентов ЛП в линейные спектральные пары (ЛСП), являющиеся математическим эквивалентом коэффициентов ЛП, но обладающие большей устойчивостью к воздействиям ошибок в канале связи;
- векторное квантование 10-мерного вектора ЛСП;
- пересчет квантованного вектора ЛСП в квантованный вектор коэффициентов синтезирующего фильтра;
- формирование текущего состояния синтезирующего фильтра на основе квантованного вектора коэффициентов.

Текущий речевой сегмент кодируется двукратной процедурой анализа через синтез: на длительности каждого полусегмента (5 мс, 40 отсчетов РС); при этом используются следующие базовые процедуры:

- формирование вариантов 40-мерного вектора $\hat{\mathbf{C}}$ сигнала возбуждения на основе двух кодовых книг и сопряженного векторного квантователя: $\hat{\mathbf{C}} = G_{\text{т}}\hat{\mathbf{C}}_{\text{т}} + G_{\text{с}}\hat{\mathbf{C}}_{\text{с}}$;
- формирование вариантов вектора синтезированной речи $\hat{\mathbf{S}}$ посредством цифровой фильтрации синтезирующим фильтром векторов $\hat{\mathbf{C}}$ сигнала возбуждения;
- расчет векторов ошибки квантования $\mathbf{e} = \mathbf{S} - \hat{\mathbf{S}}$ и их взвешивание фильтром восприятия, учитывающим особенности слухового аппарата человека;
- определение вектора $\hat{\mathbf{C}}$, обеспечивающего максимальную похожесть векторов оригинального и синтезированного РС по критерию минимума взвешенной среднеквадратической ошибки.

Ниже изложены основные особенности некоторых из указанных процедур.

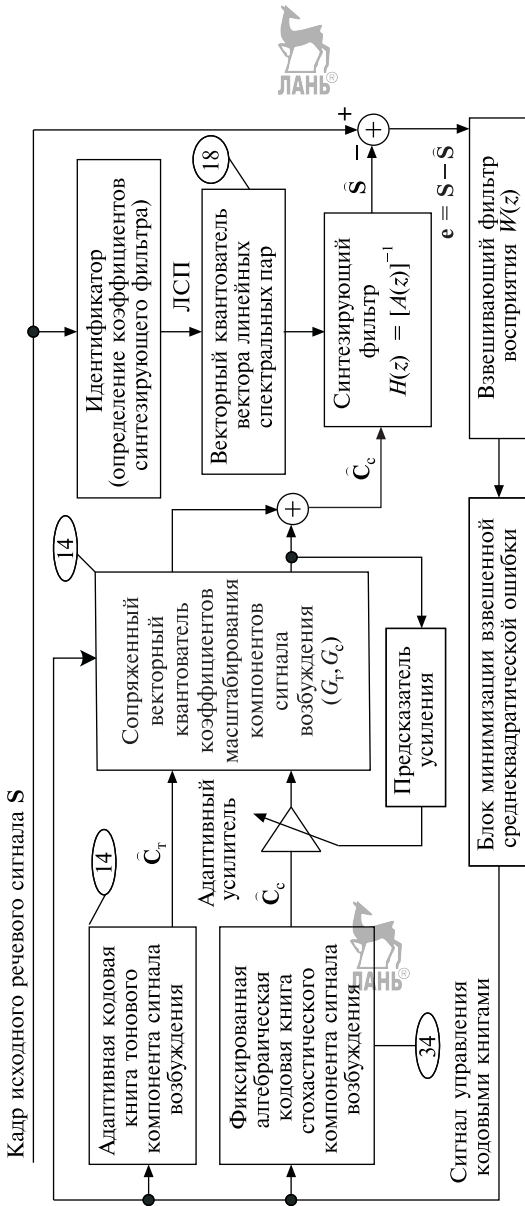


Рис. 5.27. Структурная схема кодера речевого сигнала, реализующего алгоритм G.729

Векторный квантователь вектора ЛСП использует двухкаскадное векторное квантование с межсегментовым линейным предсказанием 4-го порядка.

Тоновый компонент \hat{C}_T сигнала возбуждения отображает наличие долговременных корреляционных связей за счет вокализованной структуры речевого сигнала и несет информацию о количестве отсчетов, соответствующем периоду основного тона анализируемого сегмента речи. Соответствующая кодовая книга реализует процедуру адаптивного векторного квантования.

Стохастический компонент \hat{C}_C отображает более сложную — невокализованную — составляющую сигнала возбуждения. Соответствующая кодовая книга является фиксированной. 40-мерный вектор \hat{C}_C содержит только четыре ненулевых (± 1) символа, положение и значение которых определяются алгебраическим методом.

Сопряженный векторный квантователь содержит 4-битовую кодовую книгу коэффициентов G_T масштабирования компонента \hat{C}_T и 3-битовую кодовую книгу коэффициентов G_C масштабирования компонента \hat{C}_C с предварительным адаптивным усилением стохастического компонента сигнала возбуждения.

80-битовое кодовое слово, отображающее текущий сегмент РС и передаваемое по каналу связи от кодера к декодеру, формируется с использованием следующих информационных параметров:

- информация о текущем состоянии синтезирующего фильтра в виде номера вектора ЛСП — 18 бит;
- информация о текущем векторе \hat{C} сигнала возбуждения — 62 бита со следующим распределением информационных ресурсов:
 - 1) информация о 40-мерном векторе \hat{C}_T тонового компонента сигнала возбуждения — 13 бит (номер вектора на первом полусегменте — 8 бит, изменение вектора на втором полусегменте — 5 бит);
 - 2) информация о соответствии значений тонового компонента на предыдущем и текущем сегментах РС — 1 бит на первом полусегменте;
 - 3) номер 40-мерного вектора \hat{C}_C стохастического компонента сигнала возбуждения — 34 бита (по 17 бит на каждом полусегменте: позиции 1-го, 2-го и 3-го ненулевых символов — по 3 бита; позиция 4-го ненулевого символа — 4 бита; значения ненулевых символов — по 1 биту);
 - 4) номер 2-мерного вектора (G_T, G_C) коэффициентов масштабирования компонентов сигнала возбуждения — 14 бит (по 7 на каждом полусегменте).

Очевидно, что общая скорость кодирования РС по стандарту G.729 составляет $V_K = 80 \frac{1}{10 \cdot 10^{-3}} = 8$ кбит/с.

5.6. Кодирование широкополосного речевого сигнала

Одним из объективно возможных путей повышения качества синтезированной речи по показателям естественности ее звучания и узнаваемости абонента является расширение частотного диапазона речевого сигнала, подвергаемого обработке, с учетом в процессе кодирования психоакустических особенностей восприятия синтезированной речи человеком. В научной литературе речевой сигнал, частотный диапазон которого превышает традиционный для телефонии диапазон 0,3...3,4 кГц, принято называть широкополосным речевым сигналом (ШРС).

Объективные предпосылки повышения естественности звучания синтезированного речевого сигнала и узнаваемости абонента при переходе от традиционного телефонного РС к широкополосному не требуют доказательств в силу того, что максимальные частоты спектральных составляющих естественного речевого сигнала, формируемого артикуляционным аппаратом человека, могут превышать 15 кГц. В этих условиях искусственное ограничение частотного диапазона, изначально нацеленное на устранение частотной избыточности, приводит не только к потере энергетики РС, но и объективно ухудшает условия восприятия речи. Дополнительное включение в спектр передаваемого речевого сигнала низкочастотной области (50...300 Гц) способствует повышению натуральности синтезированной речи, комфортности и естественности телефонного диалога. В то же время кодирование высокочастотной области речевого сигнала (3400...7000 Гц) делает возможным передачу большего диапазона фрикативных различий и, следовательно, повышение разборчивости речи. Кроме того, переход к широкополосному речевому сигналу не только улучшает натуральность и разборчивость речи, но также приближает телефонный обмен к условиям естественного общения и облегчает узнаваемость абонента.

История стандартизации кодеков ШРС с диапазоном частот от 50 Гц до 7 кГц Международным союзом электросвязи началась с Рекомендации G.722 (7 kHz audio-coding within 64 kbit/s), принятой в 1988 г. Предложенный кодек использует метод полосной адаптивной дифференциальной импульсно-кодовой модуляции и формирует цифровой речевой поток со скоростями передачи 64, 56 или 48 кбит/с.

В 1999 г. МСЭ была принята Рекомендация G.722.1 (Coding at 24 and 32 kbit/s for hands-free operation in systems with low frame loss), регламентирующая трансформный речевой кодек, формирующий цифровой речевой сигнал со скоростями передачи 24 и

32 кбит/с. В качестве базисного преобразования речевого сигнала в кодеке используется одна из модификаций дискретного косинусного преобразования.

Рекомендация G.722.2 (Wideband coding of speech at around 16 kbit/s using Adaptive Multi-Rate Wideband) была принята МСЭ-Т в 2003 г., она ввела в практику адаптивный многоскоростной кодек широкополосного речевого сигнала (AMR-WB), использующий технологию кодирования речи на основе метода линейного предсказания с возбуждением от алгебраического кода (ACELP). Следует отметить, что первоначально кодек AMR-WB был разработан коллективом разработчиков телекоммуникационных систем третьего поколения (3GPP — Third Generation Partnership Project) совместно с Европейским институтом стандартов в области телекоммуникаций (ETSI — European Telecommunication Standards Institute) и стандартизован в 2001 г. для сетей мобильной связи (спецификации TS 26.190 и TS 26.194). Последующая его стандартизация МСЭ привела к тому, что кодек AMR-WB стал первым универсальным речевым кодеком, рекомендованным одновременно для проводных и беспроводных речевых служб. Важным следствием этого факта является устранение необходимости перекодировки речевого сигнала при взаимодействии речевых приложений в широком диапазоне телекоммуникационных систем.

5.6.1. Стандарт G.722

Стандарт G.722 предусматривает независимое кодирование частотных полос 0,05...4 и 4...7 кГц. Структурная схема кодера G.722 показана на рис. 5.28.

Исходный непрерывный речевой сигнал $s(t)$ после обработки аналоговым фильтром, ограничивающим спектр РС частотным диапазоном от 50 Гц до 7 кГц, поступает на аналого-цифровой преобразователь (АЦП), реализующий стандартную процедуру линейной импульсно-кодовой модуляции. С выхода АЦП, последовательно осуществляющего операции дискретизации с частотой 16 кГц, равномерного 2^{14} -уровневого квантования и двоичного примитивного кодирования, цифровой РС $x_{\text{вх}}(n)$ поступает на вход квадратурного зеркального фильтра (QMF — Quadrature Mirror Filter) передачи.

QMF-фильтр имеет один вход и два выхода, реализует операции расфильтрации цифрового РС на нижнюю и верхнюю полосы частот и передискретизации.

Использование квадратурного зеркального фильтра обусловлено тем, что его конструктивные особенности позволяют обеспечить высокое качество разделения ШРС на две полосы частот, что является существенным фактором для методов полосного кодирования.

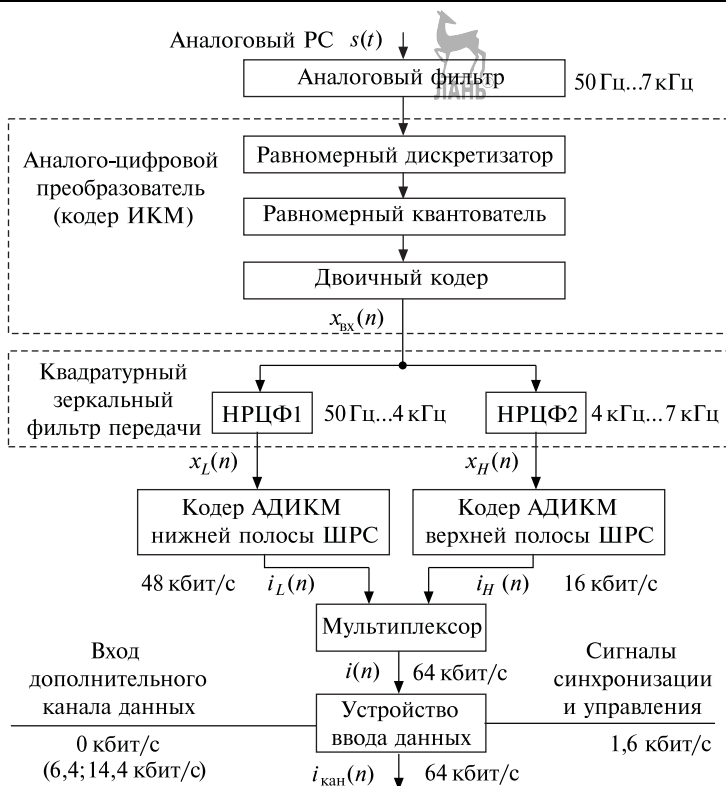


Рис. 5.28. Структурная схема кодера стандарта G.722

Функционально QMF-фильтр может быть представлен двумя параллельно включенными НРЦФ с 24-мя коэффициентами и полосами пропускания 50 Гц...4 кГц и 4 кГц...7 кГц. Оба НРЦФ имеют линейную фазочастотную характеристику. Формирование цифровых РС нижней и верхней частотных полос ШРС ($x_L(n)$ и $x_H(n)$ соответственно) сопровождается процедурой однократной децимации с коэффициентом $M = 2$, что обеспечивает переход с частоты дискретизации 16 кГц на частоту 8 кГц, используемую во всех последующих блоках кодера. Формат представления обоих сигналов составляет 15 бит. Эффективное кодирование полосовых РС осуществляется двумя независимыми кодерами АДИКМ. В обоих используются процедуры адаптивного предсказания текущего отсчета и адаптивного нелинейного квантования разности между входным и предсказанным отсчетами.

Принципиальное различие между кодерами заключается в том, что адаптивный квантователь кодера нижней полосы имеет 60 уров-

ней квантования, а адаптивный квантователь кодера верхней полосы — 4 уровня, что обуславливает формирование на периоде дискретизации 6-битных и 2-битных кодовых комбинаций сигналов $i_L(n)$ и $i_H(n)$ соответственно со структурами $I_{L1}, I_{L2}, I_{L3}, I_{L4}, I_{L5}, I_{L6}$ и I_{H1}, I_{H2} . Такое различие в процедурах квантования объясняется нахождением большей части энергии широкополосного РС в диапазоне частот до 4 кГц и большей субъективной слуховой заметностью шумов квантования в низкочастотной области.

Очевидно, что в этом случае цифровой поток $i_L(n)$ на выходе кодера АДИКМ нижней полосы РС имеет скорость 48 кбит/с, а поток $i_H(n)$ на выходе кодера верхней полосы РС — 16 кбит/с.

Мультиплексор объединяет цифровые потоки $i_L(n)$ и $i_H(n)$, формируя цифровой сигнал $i(n)$, имеющий скорость 64 кбит/с. Кодовое слово этого сигнала на периоде дискретизации представляет собой октет со структурой $I_{H1}, I_{H2}, I_{L1}, I_{L2}, I_{L3}, I_{L4}, I_{L5}, I_{L6}$.

Существенно, что стандарт G.722 предусматривает возможность передачи цифрового ШРС со скоростями 56 и 48 кбит/с. В этих режимах алгоритм кодирования речи, рассмотренный выше, остается неизменным, а скорость передачи ШРС изменяется устройством ввода данных, выполняющим замену одного (I_{L6}) или двух (I_{L5}, I_{L6}) наименее значащих битов кодового слова на выходе мультиплексора (кодовой комбинации нижней полосы ШРС) на биты цикловой синхронизации, управления режимом работы кодека и дополнительного канала данных. Скорость результирующего выходного сигнала $i_{\text{кан}}(n)$, поступающего с кодера в тракт передачи и далее на декодер, остается неизменной и во всех режимах составляет 64 кбит/с (восемь бит на период дискретизации, равный 0,125 мс).

Системные аспекты использования кодека G.722 и структура цикла цифрового сигнала $i_{\text{кан}}(n)$ дополнительно определены МСЭ-Т в рекомендациях G.725 и H.221 соответственно. В соответствии с этими документами на основе восьмого бита каждого выходного октета (позиция I_{L6}) формируется служебный канал со скоростью 8 кбит/с, который включает в себя три составляющие:

- сигнал цикловой синхронизации, имеющий скорость 0,8 кбит/с;
- сигнал распределения скоростей передачи со скоростью 0,8 кбит/с;
- прикладной канал, имеющий скорость 6,4 кбит/с, доступный для использования как в целях передачи цифрового ШРС, так и в целях организации дополнительного канала данных.

Таким образом, реальное использование кодека G.722 в системах передачи широкополосной речи предполагает три режима функционирования:

- режим 1, регламентирующий передачу ШРС со скоростью 62,4 кбит/с;
- режим 2, регламентирующий передачу ШРС со скоростью 56 кбит/с и организацию на позиции бита I_{L6} дополнительного канала данных со скоростью 6,4 кбит/с;
- режим 3, регламентирующий передачу ШРС со скоростью 48 кбит/с и организацию на позициях битов I_{L5} , I_{L6} дополнительного канала данных со скоростью 14,4 кбит/с.

Декодер стандарта G.722 реализует три различных режима декодирования, учитывающие реальную скорость передачи цифрового ШРС. Структурная схема декодера представлена на рис. 5.29.

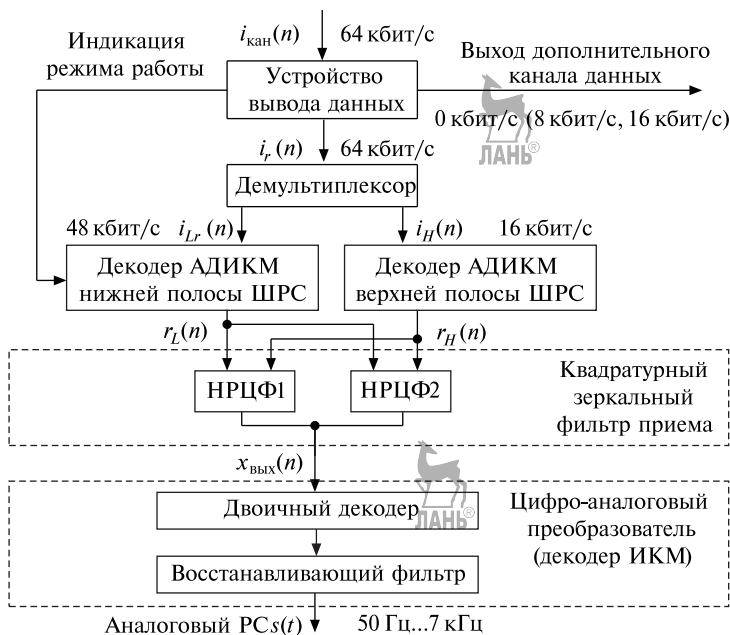


Рис. 5.29. Структурная схема декодера стандарта G.722

Устройство вывода данных в случае организации дополнительного канала данных выделяет из сигнала $i_{\text{кан}}(n)$, поступающего на декодер, биты (один или два из каждого кодового слова), соответствующие этому каналу.

Кроме того, оно формирует сигнал индикации режима работы, передаваемый на декодер АДИКМ нижней полосы ШРС. С выхода устройства вывода данных на демультимплексор поступает сигнал $i_r(n)$, скорость которого во всех режимах функционирования составляет 64 кбит/с. Битовые позиции дополнительного канала данных в

случае его организации в этом сигнале заполняются нулями. Таким образом, сигнал $i_r(n)$ в общем случае не является тождественным сигналу $i(n)$ тракта передачи. Демультимплексор разделяет входной сигнал на сигналы нижней и верхней полос ШРС. Оба сигнала имеют частоту дискретизации 8 кГц. Сигнал $i_H(n)$ верхней полосы на периоде дискретизации содержит 2 бита и в случае идеального канала связи идентичен аналогичному сигналу тракта передачи. Сигнал $i_{Lr}(n)$ на периоде дискретизации содержит 6 бит и даже в случае идеального канала связи может не соответствовать сигналу $i_L(n)$ тракта передачи из-за возможных изменений на битовых позициях дополнительного канала.

С выходов демультимплексора сигналы поступают на декодеры АДИКМ нижней и верхней полос ШРС. Результатом декодирования является формирование сигналов $r_L(n)$ и $r_H(n)$, представленных в формате 15-разрядной линейной ИКМ при частоте дискретизации 8 кГц. Режим декодирования нижней полосы ШРС (декодирование по шести, пяти или четырем старшим битам 6-битового кодового слова) выбирается с учетом сигнала индикации, поступающего на декодер от устройства вывода данных, и может быть изменен на длительности любого периода дискретизации (кодового слова).

Квадратурный зеркальный фильтр приема структурно состоит из двух НРЦФ и вспомогательных устройств. Назначением фильтра является объединение полосовых сигналов $r_L(n)$ и $r_H(n)$ с повышением частоты дискретизации выходного ШРС $x_{\text{вых}}(n)$ до 16 кГц. Отсчеты сигнала $x_{\text{вых}}(n)$ представлены в 14-разрядном формате линейной ИКМ.

Формирование восстановленного аналогового речевого сигнала с полосой частот 50...7000 Гц производится декодером импульсно-кодовой модуляции.

5.6.2. Стандарт G.722.2

В настоящее время следует считать кодек AMR-WB лучшим из реализованных решений по кодированию ШРС.

Основными характеристиками кодека AMR-WB являются:

- обработка широкополосного речевого сигнала с диапазоном частот от 50 Гц до 7 кГц;
- предварительная дискретизация ШРС с частотой 16 кГц и представление в формате 14-битной ИКМ, что формирует входной цифровой поток со скоростью передачи 224 кбит/с;
- разделение ШРС на нижнюю (50...400 Гц) и верхнюю (6400...7000 Гц) полосы частот с последующим применением отдельных процедур (трактов) их обработки (кодирования и декодирования);

Этапы и процедуры обработки и анализа речевого сигнала на длительности кадра

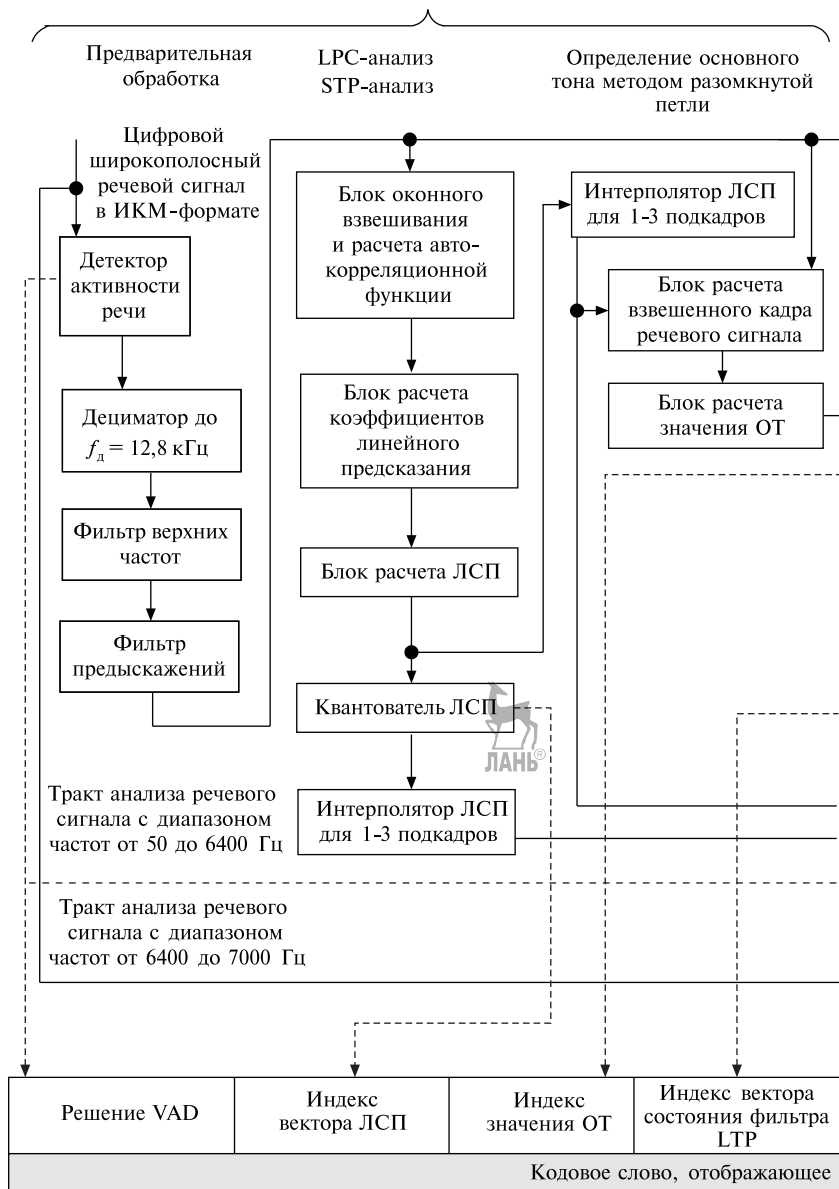


Рис. 5.30. Структурная схема кодера AMR-WB и структура

Блок
вектор
возбу



кодового слова, отображающего кадр речевого сигнала

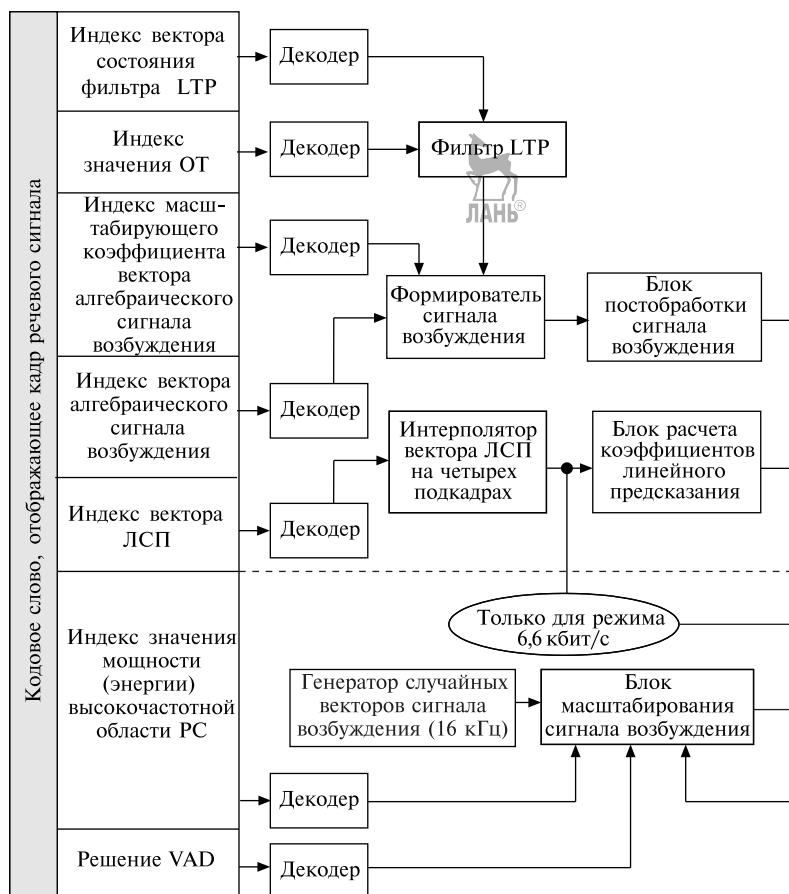
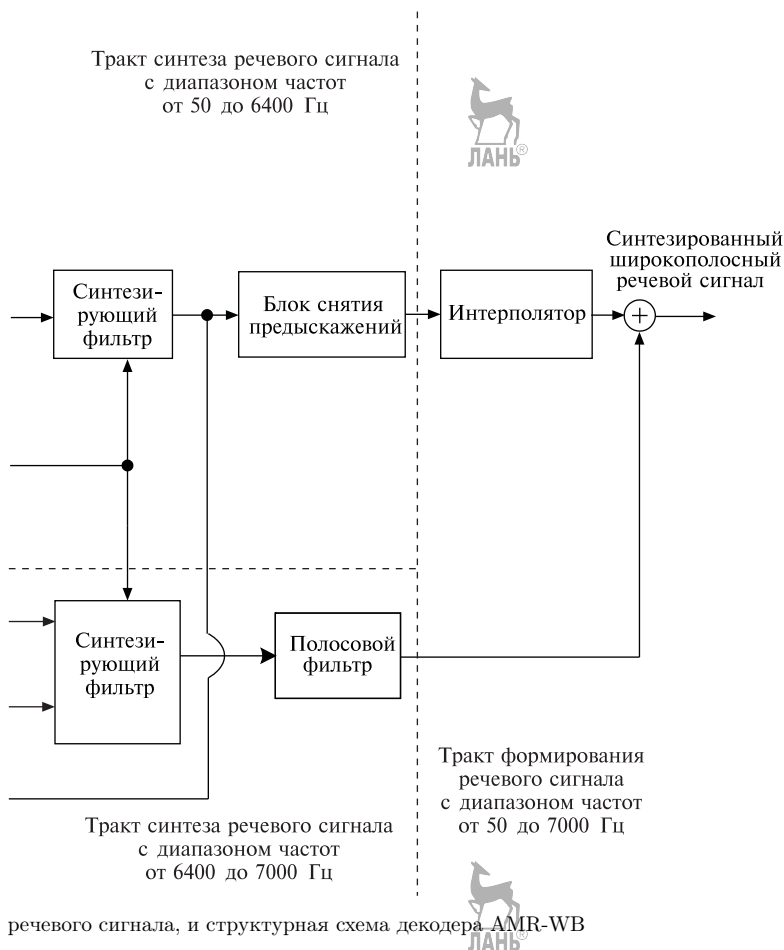


Рис. 5.31. Структура кодового слова, отображающего кадр

- использование процедуры предсказания для относительного выравнивания по мощности спектральных составляющих обрабатываемого ШРС в диапазоне частот 50...6400 Гц (обеливание речевого сигнала);
- использование традиционной локально-стационарной модели речевого сигнала с длительностью кадра в 20 мс, при этом кадр входного ИКМ-потока содержит 4480 бит, что соответствует 320 отсчетам дискретного речевого сигнала;
- использование кодирования речи на основе метода линейного предсказания с возбуждением от алгебраического кода;
- реализация процедуры кратковременного линейного предсказания (STP — Short Term Prediction) с расчетом линейных спектральных пар (ЛСП) на длительности кадра;



- реализация процедуры долговременного линейного предсказания (LTP — Long Term Prediction) на основе определения периода основного тона (ОТ);
- разделение кадра РС на четыре подкадра длительностью 5 мс для определения параметров сигнала возбуждения синтезирующего фильтра линейного предсказания;
- структурно-параметрическая степень адаптации;
- использование девяти скоростных режимов кодирования: 23,85, 23,05, 19,85, 18,25, 15,85, 14,25, 12,65, 8,85 и 6,6 кбит/с; кодовое слово, отображающее кадр речевого сигнала, составляет при этом 477, 461, 397, 365, 317, 285, 253, 177 и 132 бита;
- применение режима прерывистой передачи речи (DTX — Dis-

continuous Transmission) на основе использования детектора активности речи VAD;

- активное использование при обработке ШРС стандартных процедур цифровой обработки сигналов.

Более детальное представление о процедурах обработки ШРС в AMR-WB может быть получено на основании анализа структурных схем кодера (рис. 5.30) и декодера (рис. 5.31) и структуры кодового слова (табл. 5.7).

Таблица 5.7

Распределение битов кодового слова кодека AMR-WB по кодируемым параметрам PC для разных скоростей кодирования

Кодируемые параметры PC	Скорость кодирования речевого сигнала, кбит/с								
	6,6	8,85	12,65	14,25	15,85	18,25	19,85	23,05	23,85
Решение VAD	1	1	1	1	1	1	1	1	1
Индекс вектора ЛСП	36	46	46	46	46	46	46	46	46
Индекс вектора состояния фильтра LTP	—	—	4	4	4	4	4	4	4
Индекс значения ОТ	23	26	30	30	30	30	30	30	30
Индекс вектора алгебраического сигнала возбуждения	48	80	144	176	208	256	288	352	352
Индекс коэффициента вектора алгебраического сигнала возбуждения	24	24	28	28	28	28	28	28	28
Индекс значения мощности высокочастотной области PC	—	—	—	—	—	—	—	—	16
Общее кол-во бит	132	177	253	285	317	365	397	461	477

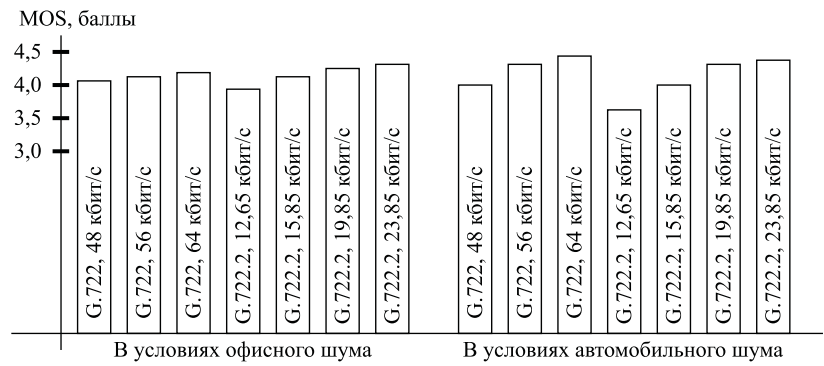


Рис. 5.32. Сравнительная оценка качества синтезированного широкополосного речевого сигнала, формируемого кодеками рекомендаций G.722 и G.722.2

На рис. 5.32 показана сравнительная оценка качества речи, обеспечиваемого кодеками G.722 и G.722.2 при работе по идеальному каналу связи на различных скоростях передачи в условиях офисного и автомобильного шумов при отношении мощностей сигнала и шума в 15 дБ. Представленные результаты одновременно свидетельствуют и о превосходстве кодека AMR-WB над известными кодеками речевого сигнала с полосой частот 0,3...3,4 кГц при сравнимых средних скоростях передачи.

Таким образом, следует считать, что технические решения по кодированию ШРС, использованные в Рекомендации G.722.2, обеспечивают значимое улучшение интегрального показателя качества синтезированной речи.

5.7. Оценка качества передачи речевого сигнала

Известно, что восприятие органами слуха человека РС является субъективным и до конца не изученным процессом, в связи, с чем точная оценка качества синтезированной речи при кодировании РС является весьма сложной и трудоемкой. Методы оценки качества синтезированной речи разделяются на субъективные и объективные.

Качество речи, воспринимаемой абонентом, может быть оценено в соответствии с тремя показателями:

- разборчивость речи (A , %);
- качество речи по показателям естественности ее звучания и узнаваемости говорящего;
- интегральное качество речи.

Преобладание интересов абонентов делает более актуальными методы субъективной оценки, характерной особенностью которых является участие в них бригад дикторов и аудиторов с определением последними качества звучания синтезированной речи. При этом ее качество определяется величиной, характеризующей субъективную оценку звучания на выходе исследуемого тракта. Основными методами субъективной оценки качества являются метод эквивалента затухания, метод мнений, метод парных сравнений и метод диагностических (артикуляционных) таблиц.

Субъективные тесты являются дорогостоящими, время- и трудоемкими, исключают возможность автоматизации процесса измерения и требуют специальной подготовки квалифицированной группы экспертов. Они обязательны при выборе стандартных кодеров, когда качество передачи речи должно быть гарантировано.

Состоятельность субъективной оценки качества речи в значительной степени зависит от выбора критерия оценки. Внедрение

новых технологий речевого анализа и синтеза и возросшие требования абонентов привели к расширению перечня критериев оценки качества: наравне с традиционными (разборчивость и громкость) все более активное применение находят естественность и узнаваемость говорящего, наличие селективных признаков искажений (картавости, гнусавости и т. п.).

Рекомендациями сектора стандартизации в области телефонии Международного союза электросвязи (МСЭ-Т) введена шкала средней субъективной оценки MOS (Mean Opinion Score) (табл. 5.8).

Шкала средней субъективной оценки MOS

Оценка	Качество	Искажения
5	Превосходно	Незначительные
4	Хорошо	Заметные, но не раздражают
3	Посредственно	Незначительно раздражают
2	Плохо	Раздражают, но не неудобны
1	Неудовлетворительно	Значительно раздражают

Кроме того, МСЭ-Т рекомендует к использованию при оценке разборчивости речи диагностический тест на рифмы DRT (Diagnostic Rhyme Test), а в качестве единицы искажения — единицу ошибки квантования QDU (Quantization Distortion Unit).

Современные гибридные кодеки по качеству синтезированной речи приближаются к кодекам формы сигнала, обеспечивая при этом меньшие скорости передачи (рис. 5.33).

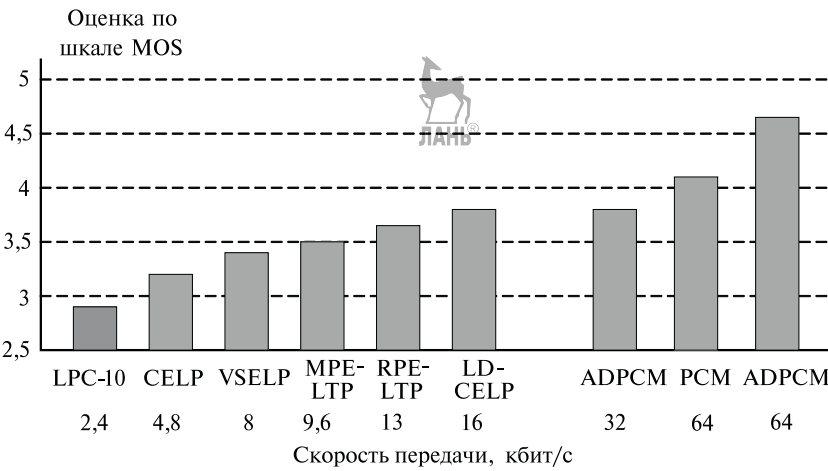


Рис. 5.33. Оценка качества речи по пятибалльной шкале MOS для различных алгоритмов кодирования

В Российской Федерации методы измерений и нормы качества передачи (воспроизведения) речи регламентированы ГОСТ Р50840-95 и Р51061-97. Данные документы устанавливают пять классов качества звучания речи на основе измерения ее разборчивости артикуляционным методом (табл. 5.9), что обусловлено высокой коррелированностью общей оценки качества звучания и степени разборчивости. В ГОСТ Р51061-97 определены нормы качества речи на основе измерения методом парных сравнений (табл. 5.10). Особенностью ГОСТ Р51061-97 является низкоскоростная направленность, что определяет его важность для диапазона скоростей < 16 кбит/с.

Таблица 5.9

Классы качества и нормы разборчивости речи

Класс качества	Характеристика класса качества	Нормы слоговой разборчивости речи, %	
		для трактов с параметрическим компандированием	для трактов с кодированием волны речевого сигнала
Высший	Понимание передаваемой речи без малейшего напряжения внимания	>93	>80
I	Понимание передаваемой речи без затруднений	86...93	56...80
II	Понимание передаваемой речи с напряжением внимания без переспросов и повторений	76...85	41...55
III	Понимание передаваемой речи с некоторым напряжением внимания, редкими переспросами и повторениями	61...75	25...40
IV	Понимание передаваемой речи с большим напряжением, частыми переспросами и повторениями	45...60	< 25

Основной подход в объективных методах оценки качества синтезированной речи — корреляционный — основан на инструментальной оценке некоторых формализованных параметров, позволяющих судить о степени похожести исходных и восстановленных на приемной стороне РС. Основными преимуществами такого подхода являются оперативность измерений, стабильность результатов и экономическая эффективность. Оценка меры искажения речи при этом может проводиться как во временной, так и в частотной областях.

Критериальный аппарат объективной оценки качества речи достаточно обширен. Наибольшую популярность приобрели среднеквадратическая и взвешенная среднеквадратическая ошибки (СКО

Таблица 5.10

Соответствие между качеством речи и оценкой в баллах

Характеристика качества речи	Баллы
Естественность звучания речи. Высокая узнаваемость. Полное отсутствие помех и искажений	4,6...5,0
Естественность звучания речи. Высокая узнаваемость. Отдельные малозаметные искажения или помехи	4,0...4,5
Естественность звучания речи. Высокая узнаваемость. Слабое постоянное присутствие отдельных видов искажений или помех	3,5...3,9
Незначительное нарушение естественности и узнаваемости. Заметное присутствие отдельных искажений или помех	3,0...3,4
Заметное нарушение естественности и ухудшение узнаваемости, присутствие нескольких видов искажений (картавость, гнусавость и др.) или помех	2,5...2,9
Существенное искажение естественности и ухудшение узнаваемости. Постоянное присутствие искажений типа картавость, гнусавость и др. или помех	1,7...2,4
Сильные искажения типа картавость, гнусавость и др. Механический голос. Наблюдается потеря естественности и узнаваемости	< 1,7

и ВСКО), мощность остатка линейного предсказания и мера Итакуры–Саито, отношение сигнал/шум (ОСШ) и сегментно-спектральная мера искажений. Эти критерии по-разному характеризуют снижение качества синтезированной речи, имеют различную вычислительную сложность, различную степень учета свойств слухового восприятия, а следовательно, требуют корректного использования в исследовательских задачах.

Они служат необходимым инструментом при проектировании и тестировании систем кодирования РС, однако для принятия правильных решений о выборе путей улучшения характеристик таких систем необходимы периодические субъективные испытания.

Наиболее общим является среднеквадратическая ошибка (MSE — Mean Square Error)

$$d(S, \hat{S}) = \frac{1}{N}(S - \hat{S})(S - \hat{S})^T,$$

где S — вектор входных значений, \hat{S} — вектор оценок.

В методе среднеквадратичной ошибки допускается, что искажения, вносимые каждым элементом вектора, имеют равный вес. Взвешенная среднеквадратичная ошибка, отражающая вклады отдельных элементов в искажение, определяется как

$$d_W(S, \hat{S}) = (S - \hat{S})(S - \hat{S})^T W,$$

где W — положительно определенная взвешивающая матрица.

В общем случае корректность критериев объективной оценки качества РС характеризуется корреляцией оценок, вычисленных с

их использованием, с оценками, полученными субъективными методами. Например, оценка качества синтезированной речи с помощью модифицированного спектрального искажения MBSD использует модель слуха человека:

$$\text{MBSD} = \frac{1}{W} \sum_{w=1}^W \sum_{n=1}^N \text{IPD}(w, n) \Delta\Lambda(w, n) [\text{сон}],$$

где $\text{IPD}(w, n)$ — индикатор перцептуального искажения громкости; $\Delta\Lambda(w, n)$ — разность между значениями громкости оригинального и искаженного w -го сегмента анализа в n -й критической частотной полосе; W — число сегментов в речевом фрагменте; N — общее количество критических частотных полос.

Согласно данному критерию синтезированный и оригинальный РС на сегменте квазистационарности подвергаются делению на критические частотные полосы, в каждой из которых вычисляются громкость сигнала и порог маскирования, далее в каждой полосе определяется разность между оригинальным и искаженным значениями громкости сигнала. Если полученное значение $\Delta\Lambda(w, n)$ превышает вычисленный порог маскирования, то индикатору перцептуального искажения громкости $\text{IPD}(w, n)$ присваивается значение 1, в противном случае значение 0.

Таблица 5.11

Субъективные и объективные оценки качества синтезированной речи

Алгоритм обработки	Скорость кодирования, кбит/с	Разборчивость речи, %	Оценка (ГОСТ 51061-97)	MBSD, мсон
CELP	2,4	85	2,8	10,93
ACELP	4,8	86	3,5	8,21
ACELP	9,6	90	3,8	4,32
ACELP	8,85	89	3,9	4,11

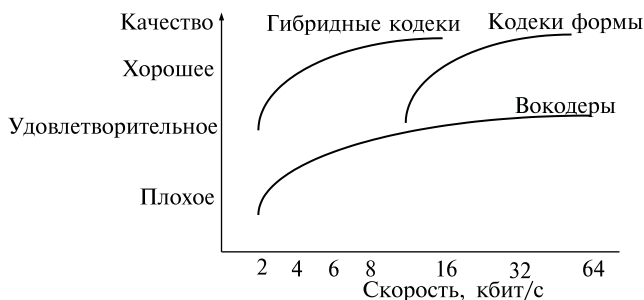


Рис. 5.34. Соотношение качества и скорости передачи

В табл. 5.11 представлены оценки качества синтезированной речи, обеспечиваемого некоторыми алгоритмами кодирования РС.

На рис. 5.34 показано соотношение качества и скорости передачи для трех основных методов кодирования речевых сигналов.



6 Цифровая обработка изображений



6.1. Математические модели изображений

При разработке и анализе систем обработки изображений удобно, а часто и необходимо иметь их математическое описание. Существуют два основных подхода к такому описанию: детерминированный и статистический. При детерминированном описании вводится математическая функция, представляющая изображение, и рассматриваются свойства изображения в каждой его точке. При статистическом описании изображение определяется усредненными характеристиками. Распределение энергии источника светового излучения по пространственным координатам x , y , времени t и длинам волн I описывается функцией $C(x, y, t, I)$. Максимальная яркость изображения ограничена возможностями источника и получателя изображения. Изображения в естественных цветах обычно характеризуются множеством $\{s_i(x, y, t)\}$ при $i = 1, 2, 3$, которое определяет красную, зеленую и синюю составляющие, черно-белого изображения в градациях серого описывается функциями $s(x, y, t)$, при этом встречаются цифровые изображения с 2, 16 и 256 уровнями серого, многоспектральные источники изображения характеризуются более широкой областью изменения параметра i (например, от 4 до 12), причем некоторые из них могут захватывать инфракрасную и ультрафиолетовую области спектра.

Источник изображения может являться как однокадровым источником отдельных независимых изображений (факсимильная связь, цифровая фотография), так и многокадровым источником коррелированных изображений (видеоинформация).

При движении объекта его математическое описание в i -м спектральном диапазоне для пространственно-временной и пространственно-спектральной областей имеет вид

$$s_i(x, y, t) = \frac{1}{(2\pi)^3} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F_i(\omega_x, \omega_y, \omega) \times$$

$$\times \exp[j(\omega_x x + \omega_y y + \omega t)] d\omega_x d\omega_y d\omega;$$

$$F_i(\omega_x, \omega_y, \omega) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s_i(x, y, t) \exp[-j(\omega_x x + \omega_y y + \omega t)] dx dy dt,$$

где ω_x, ω_y — пространственные угловые частоты, связанные с длинами волн λ_x, λ_y и с числом периодов u, v на единицу длины в направлении осей x и y соотношениями $\omega_x = 2\pi u = 2\pi/\lambda_x$ и $\omega_y = 2\pi v = 2\pi/\lambda_y$; ω — угловая частота временной оси; $F_i(\omega_x, \omega_y, \omega)$ — непрерывный пространственно-временной спектр потока изображения $s_i(x, y, t)$ в виде бесконечной суммы пространственных гармонических составляющих непрерывного аргумента.

Графическое представление потока изображений $s_i(x, y, t)$ показано на рис. 6.1, а.

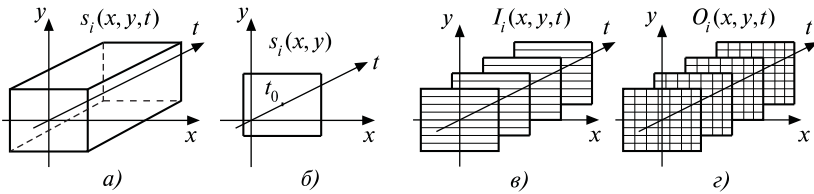


Рис. 6.1. Графическое представление моделей сигнала изображения

Неподвижное изображение не зависит от времени и является функцией координат плоскости (x, y) . Поле неподвижного изображения получается в результате временной дискретизации потока изображений $s_i(x, y, t)$, его временное и спектральное описания для любого сечения (любого t) имеют вид

$$s_i(x, y) = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F_i(\omega_x, \omega_y) \exp[j(\omega_x x + \omega_y y)] d\omega_x d\omega_y;$$

$$F_i(\omega_x, \omega_y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s_i(x, y) \exp[-j(\omega_x x + \omega_y y)] dx dy.$$

Поле неподвижного изображения, соответствующего сечению потока $s_i(x, y, t)$ в момент $t = t_0$, показано на рис. 6.1, б.

Классический метод формирования телевизионных изображений предполагает временную дискретизацию потока изображений с последующим периодическим разложением плоского поля изображения $s_i(x, y)$ на систему горизонтально расположенных строк. Такое преобразование можно в целом описать двухмерной функцией разложения $R(y, t)$, при этом формируемая передатчиком телеинформация $I_i(x, y, t)$ имеет вид $I_i(x, y, t) = s_i(x, y, t)R(y, t)$ и графически показана на рис. 6.1, в.

Применительно к цифровому телевидению функцию разложения можно представить как процесс дискретизации потока изображений в пространстве и во времени на основе трехмерной функции $D(x, y, t)$, в результате чего формируется система отсчетов $O_i(x, y, t) = s_i(x, y, t)D(x, y, t)$, представленная на рис. 6.1,2.

6.2. Описание получателя изображений

Характеристикой источника света является интенсивность излучения заданной длины волны. В зрительную систему человека свет попадает либо от самосветящегося источника, либо отраженный от некоторого предмета, либо прошедший через него.

Существуют три основные характеристики ощущения света — яркость, цветовой тон и насыщенность.

Если имеются два источника света с одинаковой формой спектральной плотности излучения, то источник с большей интенсивностью излучения воспринимается как более яркий. Известно, однако, много примеров, когда предмет с одинаковой интенсивностью излучения всех точек не воспринимается как имеющий равномерную яркость. Следовательно, интенсивность света не является адекватной количественной мерой яркости.

Признак, которым отличается, например, красный свет от зеленого, называется цветовым тоном. Если два источника света с одинаковыми спектральными плотностями наблюдать в одинаковых условиях, их цветовой тон будет одинаковым.

Третья характеристика ощущения света — его насыщенность. Этот признак позволяет отличать спектральный цвет от пастельного блеклого цвета такого же цветового тона.

Для классификации цветов удобно рассматривать их как точки некоторого цветового пространства. Но для того чтобы можно было делать количественные выводы, расстояние между двумя точками в цветовом пространстве должно соответствовать субъективно воспринимаемой разнице между представляемыми цветами, где бы ни находилась эта пара цветов.

Важным фактором для создания эффективных систем передачи изображений, предназначенных для визуального восприятия информации, является анализ особенностей зрительного анализатора человека. Очевидно, что процедура кодирования сопровождается введением в исходное изображение определенных искажений. В идеальном случае искажение изображения в процессе кодирования не должно изменять оценку восприятия исходного изображения получателем. На практике система кодирования оптимизируется по-

средством минимизации видимых искажений и учета ограничений, накладываемых зрительным анализатором человека.

Рассмотрим основные особенности зрительного анализатора человека.

Контрастная чувствительность. Известно, что реакция глаза на изменение освещения является нелинейной, контрастная чувствительность зависит от интенсивности окружающего фона. Для черно-белого изображения чувствительность наблюдателя к изменению контраста составляет около 2 %. Кроме того, установлено, что зрительный анализатор человека наиболее чувствителен к изменению в зеленом цвете, менее чувствителен к изменению в красном цвете и наименее чувствителен к изменению в синем цвете. Глаз, особенно сетчатка, имеет зрительные рецепторы двух типов: ячейки для ночного видения, воспринимающие только оттенки серого (от ярко-белого до темно-черного) и ячейки дневного видения, которые воспринимают цветовой оттенок. Первые ячейки, дающие цвет *RGB*, обнаруживают уровень яркости. Другие ячейки, ответственные за восприятие цветового оттенка, определяют величину, связанную с цветовой палитрой.

Цветовая адаптация. Воспринимаемый цветовой фон зависит от адаптации зрителя. Американский флаг, например, не сразу будет восприниматься как красно-бело-синий, если человек перед этим смотрел на интенсивный красный свет. Воспринимаемые цвета флага будут смещены в сторону голубого цвета, который является дополнительным к красному.

Цветовая слепота. Приблизительно 3...5 % всех мужчин и до 1 % женщин в той или иной форме страдают цветовой слепотой — отсутствием возможности четкого различения цветов.

Пространственная частотная характеристика зрения. Данная характеристика имеет нелинейный, в грубом приближении — полосовой характер. Таким образом, зрительная система человека нелинейна и анизотропна (не обладает инвариантностью относительно поворота). Было сделано предположение, что нелинейная реакция глаза на изменения интенсивности имеет логарифмический характер. Нелинейное преобразование осуществляется на первом этапе обработки визуальной информации в зрительной системе. Пространственно-частотные характеристики глаза объясняются частично оптическими и частично нервными механизмами. Как оптический инструмент глаз имеет ограниченную разрешающую способность из-за конечных размеров апертуры линзы, оптических аберраций и конечных размеров палочек и колбочек, что приводит к ослаблению чувствительности на высоких частотах.

Наиболее существенный вклад в частотную характеристику глаза вносит механизм латерального торможения. Нервный сигнал есть взвешенная сумма сигналов от многих соседних рецепторов. Некоторые рецепторы вносят тормозящий (т. е. отрицательный) вклад в эту сумму. Весовые множители представляют собой по существу импульсный отклик части зрительной системы, находящейся после сетчатки. Двумерное преобразование Фурье этого импульсного отклика дает соответствующую частотную характеристику.

Восприятие светового импульса, попавшего в глаз, происходит с задержкой. После окончания действия импульса ощущение света затягивается на короткое время, что вызывает пониженную чувствительность глаза человека на низких пространственных частотах.

Чувствительность глаза к искажениям понижается на контурах и перепадах цвета, что создает эффект маскирования.

6.3. Стандарты кодирования изображений

6.3.1. Стандарты кодирования неподвижных изображений

Кодирование изображений выполняет три основные функции: согласование объема алфавита источника с объемом алфавита канала, устранение избыточности источника и повышение помехоустойчивости сигнала изображения. Основа реализации представленного стандарта заключается в следующем. Низкочастотные составляющие изображений (пиксели сегмента изображения близкие по цвету) обычно гораздо более интенсивны и несут гораздо больше информации, нежели высокочастотные составляющие (пиксели сегмента изображения сильно отличающиеся по цвету), при этом низкочастотные составляющие отображают, как правило, более крупные его детали.

Стандарт JPEG (Joint Photographic Experts Group — Объединенная экспертная группа по фотографии) был разработан компанией C-Cube Microsystems (стандарт ISO/IEC 10918-1:1993 (E) Recommendation T.81 CCITT — The International Telegraph and Telephone Consultative Committee — Международный консультативный комитет по телефонии и телеграфии) как эффективный метод хранения изображений с большой глубиной цвета, получаемых, например, при сканировании фотографий с многочисленными и едва уловимыми (а иногда и неуловимыми) оттенками цвета. Формат JPEG предполагает использование алгоритмов сжатия информации с потерями и без потерь. Алгоритм сжатия без потерь так сохраняет информацию об изображении, что распакованное изображение в точности соответствует оригиналу (psx, bmp). При сжатии с потерями теряется часть информации об изображении с целью достижения большего коэффициента сжатия. Сжатие позволяет сократить необходимый для

хранения данных объем памяти, а также уменьшить время для передачи изображения или его загрузки, при этом используется принцип разрушающего сжатия. Такой подход обеспечивает высокую эффективность сжатия при приемлемом уровне потерь и допускает эффективные практические реализации. Распакованное изображение JPEG редко соответствует оригиналу абсолютно точно, но эти различия столь незначительны, что их едва можно обнаружить. Данный стандарт наиболее популярен среди методов эффективного кодирования цветных 24-битных изображений. Он является основой для разработки многих стандартов кодирования подвижных изображений.

Рассмотрим процесс сжатия изображения по стандарту JPEG. Структурная схема кодера JPEG представлена на рис. 6.2.

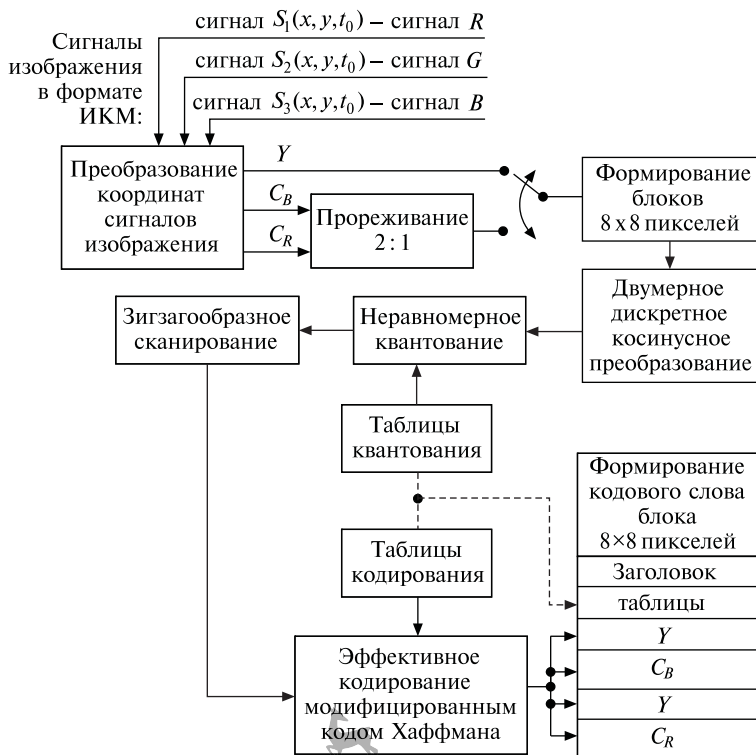


Рис. 6.2. Структурная схема кодера JPEG

Алгоритм кодирования изображения стандартом JPEG можно разделить на несколько этапов:

1. Преобразование цветового пространства.

2. Дискретное косинусное преобразование.
3. Квантование.
4. Зигзаг-сканирование.
5. Сжатие без потерь.

Этап 1. Преобразование цветового пространства.

На данном этапе выполняется преобразование координат сигналов изображений — изображение преобразуется в вид *яркость / цветность* с учетом чувствительности глаза к каждому цвету. Наибольшая степень сжатия достигается при использовании цветового пространства $YCbCr$, получаемого на основе эмпирических формул перевода:

$$\begin{aligned} Y &= 0,299R + 0,578G + 0,114B; \\ Cb &= 0,1678R - 0,3313G + 0,5B; \\ Cr &= 0,5R - 0,4187G + 0,0813B. \end{aligned} \quad (6.1)$$

В формуле (6.1) компонент Y характеризует яркость изображения и тесно связан с качеством картинки (Y — это черно-белая картинка в градациях серого). Компонент Y усредняет значения трех спектральных компонентов R, G, B (*Red* — красный, *Green* — зеленый, *Blue* — голубой цвет), значение каждого из которых может изменяться в диапазоне от 0 до 255. Компоненты Cb и Cr содержат информацию о цвете и представляют собой 2 координаты в системе, которая измеряет оттенок и насыщение цвета. Эти величины указывают на количество синего и красного в каждом пикселе, что позволяет раскрашивать Y -картинку. Эти два компонента называются *цветоразностными*. В результате формируются три плоскости значений координат (Y, Cb, Cr) цветного изображения. Все 8-битовые величины без знака (Y, Cb, Cr) в изображении — смещены по уровню: они преобразовываются в 8-битовое знаковое представление вычитанием 128 из их величины.

Для плоскостей Cb и Cr предусмотрена возможность интерполяционного прореживания (зачастую используется термин «субдискретизация») для уменьшения объема данных, подвергаемых обработке. Наиболее часто используется схема 4:2:0 (рис. 6.3).

Далее плоскости Y , Cb и Cr кодируются независимо друг от друга.

Формирование блоков 8×8 пикселей. При последующих этапах сжатия изображения, предусмотренных стандартом JPEG, в частности при дискретном косинусном преобразовании (этап 2), исходное изображение делится на небольшие неперекрывающиеся блоки изображений размером 8×8 пикселей.

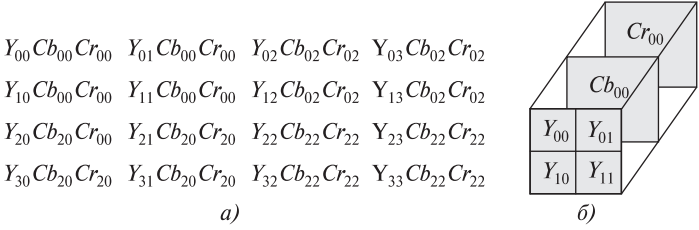


Рис. 6.3. Субдискретизация исходного блока 4×4 пиксела по схеме 4:2:0 (а) и схема 4:2:0 (б)

Размеры блока выбираются исходя из статистической взаимосвязи пикселей изображения по стандарту разложения (практически корреляционной связью между текущим пикселем и соседним можно пренебречь на расстоянии 10–12 пикселей). В алгоритме сжатия JPEG при кодировании изображения, представленного набором отсчетов, число градаций яркости составляет 256 уровней. Это самый распространенный способ хранения изображений — каждой точке на экране соответствует один байт (8 бит — 256 возможных значений), определяющий ее цвет по каждой из координат (Y, Cb, Cr) при 255 яркость максимальна, а при 0 — минимальна. Разделение исходного изображения на блоки со сторонами 8×8 пикселей представлено на рис. 6.4.

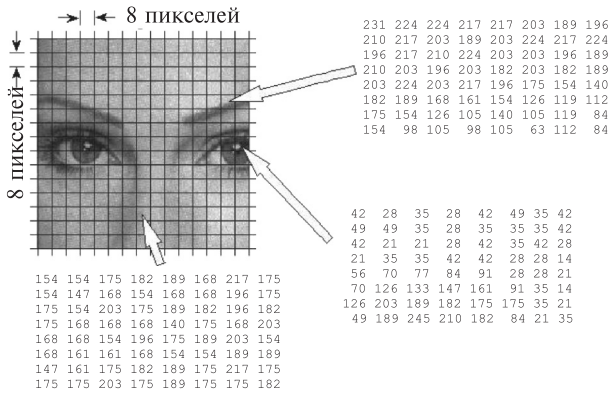


Рис. 6.4. Пример разделения исходного изображения на блоки со сторонами 8×8 пикселей для компонента Y

На последующих этапах производится сжатие каждого блока отдельно. Так как исходное изображение разделяется на квадратные блоки по 8×8 пикселей, то обязательным требованием стандарта является кратность размера изображения восьми пикселям как по высоте, так и по ширине (по координатам *x* и *y*).

Этап 2. Дискретное косинусное преобразование (ДКП).

Цель такого преобразования состоит в том, чтобы вместо обработки значений (Y, Cb, Cr) цветного изображения работать с пространством параметров их спектральных составляющих: изменения яркости и оттенка, которые тесно связаны с уровнем детализации изображения, причем более высокие значения частот составляющих соответствуют высокому уровню детализации.

Математическое определение прямого ДКП (FDCT — Forward discrete cosine transform) и обратного ДКП (IDCT — inverse discrete cosine transform), используемых в стандарте JPEG:

$$F(u, v) = C(u, v) \sum_{x=0}^7 \sum_{y=0}^7 s(x, y) \cos \left(\frac{2x+1}{16} u \pi \right) \left(\frac{2y+1}{16} v \pi \right);$$

$$s(x, y) = \frac{1}{4} \sum_{x=0}^7 \sum_{y=0}^7 C(u, v) F(u, v) \cos \left(\frac{2x+1}{16} u \pi \right) \left(\frac{2y+1}{16} v \pi \right),$$

где $s(x, y)$ — матрица значений любой из координат (Y, Cb, Cr) для блока 8×8 пикселей; $F(u, v)$ — значения, полученные в результате преобразования (амплитуды частотного пространства); (x, y) — координаты отдельной точки изображения блока 8×8 ; (u, v) — координаты спектральной составляющей двумерного разложения.

Сомножитель $C(u, v)$ определяется следующим соотношением:

$$C(u, v) = \begin{cases} 1/\sqrt{2}, & u = v = 0; \\ 1, & u, v = 1, 2, \dots, N-1. \end{cases}$$

С помощью двумерного ДКП преобразуются блоки 8×8 пикселей матриц (Y, Cb, Cr) с информацией о яркости Y и цветоразности Cb и Cr пикселей в матрицу такого же по размерам блока 8×8 коэффициентов ДКП, статистические связи в котором в значительной степени ослаблены. Данные коэффициенты трансформанты подразделяются на один DC-коэффициент трансформанты (характеризующий амплитуду постоянной составляющей) двумерного преобразования и на 63 AC-коэффициента трансформанты (характеризующих амплитуды переменных составляющих двумерного преобразования), отражающих различные спектральные составляющие в изображении.

Наибольшие значения коэффициентов ДКП сосредоточены в левом верхнем углу возле DC-коэффициента (низкочастотные составляющие). Его правая нижняя часть (высокочастотные составляющие) заполнена относительно небольшими числами. Таким образом, при сжатии изображения большая часть коэффициентов ДКП имеет равное нулю или близкое к нему значение, что позволяет исключить данные коэффициенты из передачи.

Важным свойством ДКП является обратимость его процедуры. Другими словами, применение процедуры обратного ДКП к блокам коэффициентов ДКП позволяет получить исходный блок изображения. В случае моноцветного изображения только один коэффициент ДС имеет ненулевое значение. Обычно исходное изображение не такое простое, и АС-коэффициенты не равны нулю. Сжатие данных при ДКП связано с экономным квантованием коэффициентов ДКП, т. е. при разложении матрицы области изображения в двойной ряд по косинусам оказываются значимыми только первые коэффициенты. Таким образом, сжатие с потерями в JPEG осуществляется за счет устранения (обнуления) малых значений амплитудного спектра в реальных изображениях.

Этап 3. Квантование.

На этом этапе коэффициенты ДКП делятся на некоторое число в соответствии с матрицей квантования, а полученное значение округляется до целого. Матрица квантования опирается на «психовизуальный порог», ее применение приводит к хорошему компромиссу сжатия и качества для изображений с 8-битовой яркостью и оттенками. Она может быть фиксированной или (для более качественного и эффективного сжатия) получена в результате анализа исходной картинки. Также имеется возможность изменения значений матрицы квантования на основе компромисса между размером конечного закодированного изображения и качеством его восприятия. В табл. 6.1 представлен результат выполнения процедуры квантования над полученной ранее матрицей коэффициентов ДКП при использовании стандартной матрицы квантования.

В результате квантования удаляются параметры спектральных составляющих высоких частот, которые отвечают за высокую детализацию. Глаз более чувствителен к параметрам спектральных составляющих низких частот, чем высоких, таким образом, их удаление практически не влияет на визуальное восприятие изображения.

От выбора таблицы квантования в значительной степени зависит как эффективность сжатия — число нулей в квантованном (округленном) спектре, так и качество восстановленной картинки. АС-коэффициенты обычно подвергают линейному квантованию с учетом «квантующей матрицы», которая каждому из АС-коэффициентов ставит в однозначное соответствие шаг линейного квантователя (чем выше в частотном аспекте номер АС-коэффициента, тем больший шаг квантования можно выбрать). На практике это приводит к поэлементному делению полученной матрицы ДКП размером 8×8 пикселей на значения коэффициентов матрицы квантования с последующим округлением полученных результатов до целых зна-

Таблица 6.1

Процедура квантования в стандарте JPEG

Матрица коэффициентов ДКП							
−288,6	627,11	−4,1	−5,2	2,1	−1,7	−2,7	1,3
539,87	−5,5	322,2	23,1	−2,9	−0,1	0,4	−1,2
−6,9	−9,3	−1,6	1,5	0,2	−0,9	−0,6	−0,1
−7,1	−1,9	0,2	1,5	0,9	−0,1	0	0,3
−420,21	−0,8	1,5	1,6	−0,1	−0,7	0,6	1,3
−288,3	−0,2	1,6	−0,3	−0,8	1,5	1	−1
−1,3	−0,4	−0,3	−1,5	−0,5	1,7	1,1	−0,8
−2,6	1,6	−3,8	−1,8	1,9	1,2	−0,6	−0,4
Матрица квантования							
16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99
Результат квантования							
−18	57	0	0	0	0	0	0
45	0	23	1	0	0	0	0
0	0	0	0	0	0	0	0
−30	0	0	0	0	0	0	0
−16	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

чений (дробная часть отсутствует). Таким образом, формируется целочисленная матрица размером 8×8 квантованных значений коэффициентов ДКП.

Впоследствии на квантованном векторе, на высоких пространственных частотах, формируются длинные последовательности нулей. В результате на представление каждого из АС-коэффициентов тратится меньше бит, чем на ДС-коэффициент, и меньшее количество бит, чем для каждого из элементов исходного кодируемого блока изображения. После квантования многие АС-коэффициенты обращаются в нуль, при этом ненулевые коэффициенты чаще всего концентрируются вблизи ДС-коэффициента.

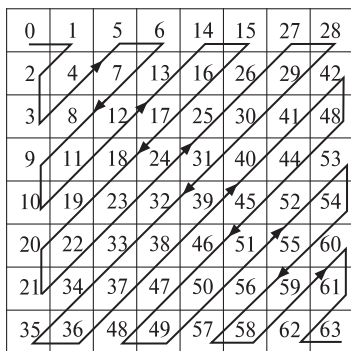
При сжатии изображения по стандарту JPEG потери информации происходят на этапе квантования, его устранение приводит к разновидности стандарта JPEG без потерь. Чем больше значения в матрице квантования, тем больше отбрасывается информации из изображения и тем более плотно сжимается изображение. Компромисс заключается в том, что более высокие значения матрицы квантования приводят к худшему качеству изображения. При формиро-

вании изображения JPEG пользователь устанавливает показатель качества, который «управляет» значениями матрицы квантования. Оптимальные значения, обеспечивающие лучший баланс между коэффициентом сжатия и качеством изображения для разных изображений, могут быть найдены только подбором.

Этап 4. Зигзаг-сканирование.

Применение зигзаг-сканирования квантованных коэффициентов ДКП позволяет экономно закодировать длины последовательно встречающихся нулевых АС-коэффициентов. Процедура зигзаг-сканирования представлена на рис. 6.5.

Поскольку в квантованной матрице отсутствует значительная доля высокочастотной информации, имеющейся в исходной матрице, то ее представление потребует в среднем до половины своего первоначального размера и даже менее. Реальные фотографические изображения часто совсем невозможно сжать с помощью методов сжатия без потерь, поэтому 50%-ное сжатие можно признать достаточно хорошим. С другой стороны, применяя методы сжатия без потерь, можно сжимать некоторые изображения, например чертежи, на 90 % однако такие изображения плохо подходят для сжатия методом JPEG.



0	1	5	6	14	15	27	28
2	4	7	13	16	26	29	42
3	8	12	17	25	30	41	48
9	11	18	24	31	40	44	53
10	19	23	32	39	45	52	54
20	22	33	38	46	51	55	60
21	34	37	47	50	56	59	61
35	36	48	49	57	58	62	63

Рис. 6.5. Процедура зигзаг-сканирования

но, что для этого нужно прочесть элементы матрицы коэффициентов ДКП в порядке, изображенном на рис. 6.5, т. е. зигзагообразно — из левого верхнего угла к правому нижнему. Эта процедура и называется зигзаг-сканированием.

После считывания по зигзагу матрицы 8×8 формируется вектор с 64 коэффициентами (0–63), смысл которого состоит в том, что коэффициенты ДКП считываются в порядке повышения пространственных частот. В итоге формируется вектор, отсортированный критериями пространственной частоты: первая элемент (индекс 0)

После того как выполнено ДКП-преобразование над блоком 8×8 и выполнено его квантование, формируется новый блок 8×8 , который просматривается по зигзагу (рис. 6.5, числа в блоке 8×8 указывают порядок считывания квантованных коэффициентов).

Зигзаг-сканирование осуществляется таким образом, чтобы сгруппировать по возможности вместе все большие значения и все нулевые значения спектра. Совершенно очевидно,

соответствует самой низкой частоте в изображении (квантованному коэффициенту DC), а с увеличением индекса вектора позиционируются элементы, соответствующие высшим частотам (элемент с индексом 63 соответствует амплитуде самой высокой частоты в блоке 8×8).

В результате такого преобразования квадратная матрица 8×8 квантованных коэффициентов ДКП преобразуется в линейную последовательность из 64 чисел, большая часть из которых — это идущие подряд нули. Известно, что такие потоки можно эффективно сжимать путем кодирования длин повторений.

DC-коэффициенты содержат большую часть энергии, присутствующей в исходном изображении блока 8×8 , при этом имеется очень близкое соотношение между коэффициентом DC последовательных блоков. Данная особенность определяет порядок представления DC-коэффициентов соседних блоков 8×8 изображения — они кодируются дифференциально. Так, кодированию подвергается различие между DC последовательных блоков 8×8 , а уже затем данная величина подвергается сжатию без потерь.

Рассмотрим исходный вектор AC-коэффициентов длиной 63 элемента (отсутствует DC-коэффициент).

Допустим, имеется следующая последовательность AC-коэффициентов: $\{57, 45, 0, 0, 0, 0, 23, 0, -30, -16, 0, 0, 1, 0, 0, \dots, 0\}$.

Для рассмотренного примера получим:

$(0, 57); (0, 45); (4, 23); (1, -30); (0, -16); (2, 1);$ ЕОВ.

Представление коэффициента AC, отличного от нуля, осуществляется путем записи пары значений, первое из которых — число последовательных предшествующих нулей перед значащими цифрами, а второе — сам AC-коэффициент, отличный от нуля; если после значащего AC-коэффициента все остальные AC-коэффициенты равны нулю, то ставится маркер ЕОВ (End of block). ЕОВ — короткая форма, определяющая окончание обхода блока — специальная кодированная величина (маркер 1010).

При обходе число предшествующих нулей представляется 4-битовой величиной и, следовательно, не может превысить 15. Если же при обходе число предшествующих нулей превышает значение 15, то ставится специальная пара значений (15, 0) — специальная кодированная величина, которая указывает на факт присутствия при обходе 16 последовательных нулей, именно 16, а не 15, так как 15 нулей представляется обычным способом. После записи в последовательность пары значений (15, 0) подсчет нулей до ненулевого AC-коэффициента начинается заново.

Этап 5. Сжатие без потерь.

определило особенность используемого кодирования на данном этапе. Вместо хранения фактической величины стандарт JPEG определяет минимальный размер в битах, в котором мы можем представить эту величину и который назван категорией. Именно категория величины подвергается кодированию по Хаффману с использованием префиксного кода, а следующее, точно известное количество бит, определяет значение соответствующего параметра.

DC-коэффициент соответствует постоянной составляющей блока изображения подобно средней величине. (Примечание: последовательные блоки 8×8 одинаковых компонентов изображения, подобно последовательным блокам 8×8 для Y , или последовательные блоки для Cb , или для Cr .) Так как DC-коэффициенты предварительно подвергаются дифференциальному представлению, то большинство из них малы, что определяет малое количество битов на кодирование наиболее вероятных категорий хранения значения этих величин. В стандарте JPEG вводит 16 категорий, номер категории определяет количество битов для представления соответствующих величин внутри каждой из них.

При кодировании дифференциального значения DC-коэффициента величиной, например, (-18) , согласно табл. 6.2, он попадает в 5-ю категорию, таким образом, на его представление необходимо будет задействовать 5 бит. Номер категории будет закодирован кодом Хаффмана, а сам коэффициент будет представлен двоичной последовательностью 01101, что соответствует порядковому номеру 13 в диапазоне возможных значений изменения коэффициента, попавшего в пятую категорию (табл. 6.3).

При кодировании AC-коэффициентов, представленных двумерным вектором, например вектора $(4, 23)$, изначально подход идентичен представленному выше для DC-коэффициента. Так, значение 23 попадает в пятую категорию и кодируется в ней двоичным кодом в виде 10111. Отличие от представления DC-коэффициента состоит в том, что префиксному кодированию Хаффмана подвергается сформированный байт, первые четыре бита которого определяют количество предшествующих нулей до ненулевого AC-коэффициента (в данном случае 4), а вторые четыре бита — категорию значения AC-коэффициента (в данном случае 5). Таким образом, сформированный байт представляет собой последовательность 01000101 или в десятичной системе счисления 69, он и подвергается префиксному кодированию Хаффмана.

Кодирование Хаффмана происходит по заданной стандартной таблице Хаффмана, которая определяет сам код и для которой протестировано большинство изображений с хорошими показателями

Таблица 6.3

Порядок определения категории кодом Хаффмана

Величины	Категория	Возможные биты для представления величины в категории
0	0	
1,1	1	0,1
-3,-2,2,3	2	00, 01, 10, 11
-7,...,-4,4,...,7	3	000, 001, 010,011,100,101,110,111
-15,...,-8,8,...,15	4	0000,...0111,1000...1111
-31,...,-16,16,...,31	5	00000...01111,10000...11111
-63,...,-32,32,...,63	6	000000...011111,100000...111111
-127,...,-64,64,...,127	7	0000000...011111,1000000...1111111
-255,...,-128,128,...,255	8	00000000...11111111
-511,...,-256,256,...,511	9	000000000...111111111
-1023,...,-512,512,...,1023	10	0000000000...1111111111
-2047,...,-1024,1024,...,2047	11	00000000000...11111111111
-4095,...,-2048,2048,...,4095	12	000000000000...111111111111
-8191,...,-4096,4096,...,8191	13	0000000000000...1111111111111
-16383,...,-8192,8192,...,16383	14	00000000000000...11111111111111
-32767,...,-6384,16384,...,32767	15	000000000000000...111111111111111

сжатия. Таким образом, кодирование Хаффмана представляет собой неразрушающее эффективное статистическое кодирование, а с учетом характерной структуры данных, подвергаемых обработке, оно существенно уменьшает их объем.

По завершении данного этапа происходит структуризация выходной последовательности, которая совместно с маркерами и битами данных, отвечающими за параметры сжатия, формирует окончательное сжатое представление изображения.

JPEG 2000. Стандарт JPEG 2000 поддерживает более эффективные методы сжатия по сравнению с JPEG и основан на использовании двумерного вейвлет преобразования, составляющего основу данного стандарта кодирования неподвижных изображений.

Основные отличия стандарта сжатия изображений JPEG 2000 от JPEG заключаются в следующем:

- вместо дискретного косинусного преобразования используется многоуровневое дискретное вейвлет-преобразование (DWT — discrete wavelet transform);
- вместо кодирования по Хаффману используется арифметическое сжатие;
- в стандарт изначально заложено управление качеством областей изображения;

Последовательность операций в стандарте JPEG 2000 представлена на рис. 6.7.

Этап 1. В JPEG 2000 предусмотрен сдвиг яркости (DC level shift) каждого компонента (RGB) изображения перед преобразова-

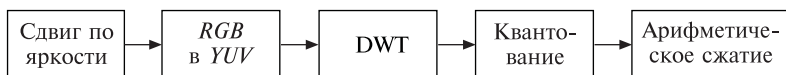


Рис. 6.7. Последовательность операций, используемых в стандарте JPEG 2000

нием в YUV . Это делается для выравнивания динамического диапазона, что приводит к увеличению степени сжатия. Данное преобразование определяется следующим выражением:

$$I'(x, y) = I(x, y) - 2^{ST-1},$$

где $I(x, y)$ — исходное изображение; $I'(x, y)$ — изображение, сдвинутое по яркости.

Значение степени ST для каждого компонента R , G и B определяется при сжатии компрессором. При восстановлении изображения выполняется обратное преобразование:

$$I'(x, y) = I(x, y) - 2^{ST+1}.$$

Этап 2. Перевод изображения из цветового пространства RGB с компонентами, отвечающими за красную (*Red*), зеленую (*Green*) и синюю (*Blue*) составляющие цвета точки, в цветовое пространство $YCbCr$. Формулы перевода определяются в соответствии с выражением (6.1) аналогично стандарту JPEG. Также возможно построение другой цветовой схемы YUV , являющейся обратимой процедурой и применяемой при сжатии без потерь.

$$\begin{vmatrix} Y \\ U \\ V \end{vmatrix} \begin{pmatrix} (R + 2G + B)/4 \\ R - G \\ B - G \end{pmatrix}.$$

Обратное преобразование осуществляется с помощью обратной матрицы:

$$\begin{vmatrix} R \\ G \\ B \end{vmatrix} \begin{pmatrix} U + G \\ Y - (U + V)/4 \\ V + G \end{pmatrix}.$$

Такое представление изображения дает возможность повысить коэффициент сжатия, так как компоненты Cb , Cr либо UV имеют большинство значений, сгруппированных возле нуля.

Этап 3. После преобразования цветового пространства каждый сегмент изображения подвергается процедуре многоуровневого дискретного вейвлет-преобразования. Для этого используют низкочастотные $h_{нч}(n)$ и высокочастотные $h_{вч}(n)$ вейвлет-фильтры, относящиеся к классу биортогональных фильтров с КИХ.

Для преобразования могут использоваться различные вейвлет-фильтры. Обязательная часть стандарта предписывает использование только двух фильтров для сжатия с потерями и без потерь, их

коэффициенты представлены в табл. 6.4. Однако расширение стандарта JPEG 2000 допускает и любые другие фильтры. Дискретное wavelet-преобразование может быть двух видов — для случая сжатия с потерями и для сжатия без потерь.

Таблица 6.4

Коэффициенты КИХ-фильтров для сжатия с потерями и без потерь

i	Низкочастотные коэффициенты	Высокочастотные коэффициенты	
Коэффициенты при упаковке			
0	1,115087052456994	0,6029490182363579	
1	0,591271763114247	−0,2668641184428723	
2	−0,05754352622849957	−0,07822326652898785	
3	−0,09127176311424948	0,01686411844287495	
4	0	0,02674855741080976	
другие i	0	0	
Коэффициенты при распаковке			
0	0,6029490182363579	1,115087052456994	
1	−0,2668641184428723	0,591271763114247	
2	−0,0782236652898785	−0,05754352622849957	
3	0,01686411844287495	−0,09127176311424948	
4	0,02674855741080976	0	
другие i	0	0	
Коэффициенты для сжатия без потерь			
При упаковке		При распаковке	
Низкочастотные коэффициенты	Высокочастотные коэффициенты	Низкочастотные коэффициенты	Высокочастотные коэффициенты
6/8	1	1	6/8
2/8	−1/2	1/2	−2/8
−1/8	0	0	−1/8

Первоначально изображение подвергается чередующимся последовательностям вертикальных и горизонтальных одномерных вейвлет-преобразований. Сначала преобразуются все строки, а затем все столбцы. Результатом разложения являются 4 матрицы: HH_1 , HL_1 , LH_1 , LL_1 , соответствующие фильтрации фильтром $h_{ВЧ}(n)(H)$ по строкам и по столбцам, фильтром $h_{ВЧ}(n)(H)$ по строкам и $h_{НЧ}(n)(L)$ по столбцам, фильтром $h_{НЧ}(n)(L)$ по строкам и $h_{ВЧ}(n)(H)$ по столбцам, фильтром $h_{НЧ}(n)$ по строкам и столбцам (рис. 6.8). Далее производится децимация полученных матриц по строкам и столбцам с коэффициентом 2. По сути данная процедура представляет собой алгоритм кратномасштабного вейвлет-анализа, представленного в п. 1.17.3 и проиллюстрированного на рис. 1.41.

В результате такого преобразования получаются низкочастотная составляющая (уменьшенная копия изображения), а также информация о вертикальных, горизонтальных и диагональных высо-

кочастотных элементах изображения. При таком преобразовании общее число коэффициентов преобразования равно числу пикселей исходного уровня LL_0 , однако наиболее значимая для восприятия информация находится именно в области LL_1 . Затем матрица LL_1 подвергается дальнейшему вейвлет-преобразованию.

Результатом вейвлет-преобразования являются матрицы HH_2 , HL_2 , LH_2 , LL_2 ; в LL_2 масштаб исходного изображения уменьшен уже в 4 раза. В общей рекомендации стандарт не определяет число разложений LL_0 . Такая процедура может повторяться несколько раз. Результатом преобразования на каждом уровне разложения является набор из четырех матриц уменьшающейся размерности. На следующем этапе матрица LL_i , получившаяся в результате предыдущего преобразования, снова преобразуется (сначала все строки, затем все столбцы) и так далее. Такая схема преобразования получила название «лифтинг-схема», которая в качестве примера приведена на рис. 6.8. Каждая из матриц разложения делится на блоки, которые затем кодируются независимо друг от друга с помощью контекстного адаптивного двоичного арифметического кодера. В результате преобразования получают множество прямоугольных диапазонов вейвлет-коэффициентов, которые принято называть частотными диапазонами, так как они содержат информацию о том, как ведет себя исходный двумерный сигнал (изображение) при разном разрешении (т. е. набор коэффициентов при разной частоте).

Само преобразование в одномерном случае представляет собой скалярное произведение коэффициентов фильтра на строку преобразуемых значений (в данном случае — на строку изображения). При этом четные выходящие значения формируются с помощью низкочастотного преобразования, а нечетные значения — с помощью высокочастотного преобразования в соответствии с выражениями:

$$y_{\text{output}}(2n) = \sum_{j=0}^{N-1} x_{\text{input}}(j) h_{\text{НЧ}}(j - 2n);$$

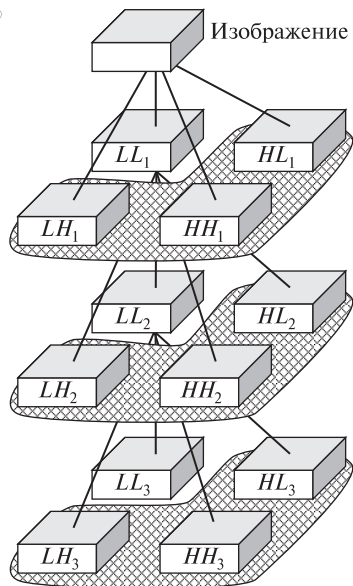


Рис. 6.8. Лифтинг-схема вейвлет-преобразования

$$y_{\text{output}}(2n + 1) = \sum_{j=0}^{N-1} x_{\text{input}}(j)h_{\text{ВЧ}}(j - 2n - 1),$$

Далее к строке применяется чересстрочное преобразование, при котором все четные коэффициенты переписываются в начало строки, а все нечетные — в конец (табл. 6.5). При таком преобразовании в начале строки формируется «уменьшенная копия» всей строки (низкочастотная составляющая), а в конце строки — информация о колебаниях значений промежуточных пикселей (высокочастотная составляющая).

Таблица 6.5

Процедура чересстрочного преобразования

y_{output}	1	0	3	1	11	4	13	−2	8	−5
y'_{output}	1	3	11	13	8	0	1	4	−2	−5

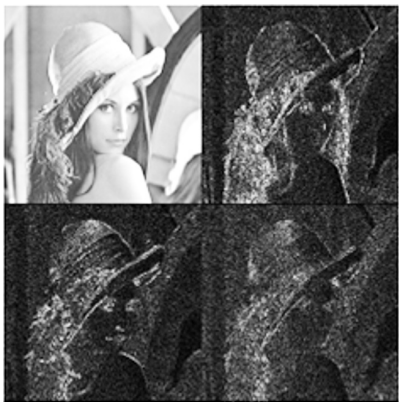


Рис. 6.9. Однократное применение двумерного вейвлет-преобразования

Таким образом, преобразование применяется сначала ко всем строкам изображения, а затем ко всем его столбцам. В результате изображение делится на 4 квадранта (рис. 6.9). В первом квадранте формируется уменьшенная копия изображения, а в остальных трех — высокочастотная информация, как было указано выше. Преобразование повторно применяется уже только к первому квадранту изображения по тем же правилам (преобразование второго уровня) и т. д.

Для сохранения результатов под данные 2-го и 3-го квадрантов выделяется на 1 бит больше, а под данные 4-го квадранта — на 2 бита больше. Так, если исходные данные были 8-битовые, то на 2-й и 3-й квадранты нужно 9 бит, а на 4-й — 10, независимо от уровня применения DWT. При записи коэффициентов в файл можно использовать иерархическую структуру DWT, помещая коэффициенты преобразований с большего уровня в начало файла. Это позволяет получать «изображение для предварительного просмотра», прочитав небольшой участок данных из начала файла, и не распаковывать весь файл, как это приходилось делать при сжатии изображения целиком в стандарте JPEG. Иерархичность преобразования

может использоваться также для плавного улучшения качества изображения при его передаче.

Этап 4. Аналогично стандарту JPEG в JPEG 2000 после DWT выполняется процесс квантования.

Коэффициенты соответствующих уровней делятся на заранее заданное число. При увеличении этого числа снижается динамический диапазон коэффициентов, они становятся ближе к нулю и получают большую степень сжатия. Варьируя эти числа для разных уровней преобразования, для разных цветовых компонентов и разных квадрантов, можно управлять степенью потерь в изображении. Рассчитанные в компрессоре оптимальные коэффициенты квантования передаются в декомпрессор для однозначной распаковки.

Именно на этапе квантования возникают основные информационные потери, но за счет квантования возможно существенное уменьшение объема представления изображения. Если производится сжатие без потерь, в квантовании нет необходимости — оно может быть либо равномерным скалярным, либо каким-либо другим (например, векторным). В случае использования равномерного скалярного квантования квант-параметр может изменяться в зависимости от квантуемого диапазона.

Этап 5. Этап арифметического кодирования.

Вейвлет-коэффициенты обрабатываются методом контекстно-зависимого бит-ориентированного арифметического кодирования. Для сжатия получающихся массивов данных в JPEG 2000 используется вариант арифметического сжатия.

Этап арифметического кодирования является завершающим этапом кодирования и осуществляется на основе алгоритма, предложенного Дэвидом Таубманом. Диапазоны коэффициентов разделяются на прямоугольные кодовые блоки (как правило, 32×32 или 64×64), что иллюстрируется на рис. 6.10.

Каждый из блоков кодируется независимо, образуя элементарную кодовую последовательность. Это означает, что состояние арифметического кодера

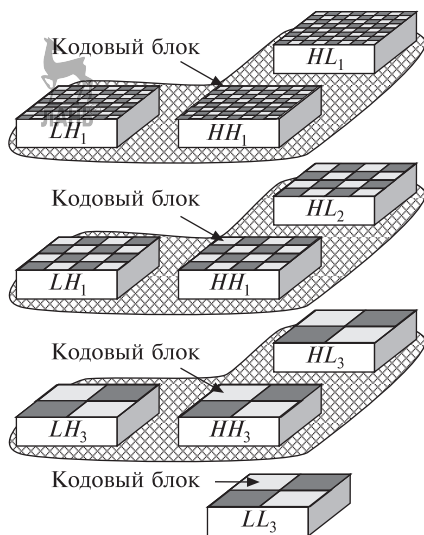


Рис. 6.10. Разделение на прямоугольные кодовые блоки

состояние арифметического кодера

сбрасывается перед кодированием очередного кодового блока. Полный сжатый кодовый поток построен на основе комбинирования кодовых потоков отдельных блоков.

В процессе кодирования коэффициенты в блоке виртуально представляются в виде битовых плоскостей. При кодировании коэффициентов матриц *вейвлет-коэффициентов* в блоке виртуально представляются в двоичном виде. Все биты коэффициентов, относящихся к одному разряду, образуют *битовую плоскость*. Кодирование коэффициентов сводится к кодированию битов, составляющих эти коэффициенты.

Таким образом, арифметическое кодирование является бит-ориентированным. При этом каждая битовая плоскость кодируется с помощью арифметического кодера за три прохода: первый называется *significance propagation* (распространение значащего разряда), второй — *magnitude refinement* (уточнение значения) и третий — *cleanup* (доочистка).

Первую плоскость составляют знаки коэффициентов, остальные плоскости соответствуют различным разрядам коэффициентов (положение бита в плоскости соответствует положению коэффициента в блоке). Кодирование коэффициентов сводится к кодированию битов, составляющих эти коэффициенты. Таким образом, арифметическое кодирование является бит-ориентированным.

Арифметическое кодирование основано на контекстно-зависимой модели. Контекст формируется как функция от значений битов, окружающих кодируемый бит. Кодирование реализуется по плоскостям: сначала кодируется плоскость, соответствующая старшему разряду коэффициентов, затем следующая по убыванию и т. д. Во время кодирования каждому коэффициенту в кодируемом блоке ставится в соответствие параметр *значимость*, соответственно 0 или 1. Коэффициент называется значимым, если в уже закодированных на данный момент битовых плоскостях присутствует хотя бы один ненулевой разряд данного коэффициента. Каждая битовая плоскость кодируется в три прохода.

Во время первого кодового прохода распространяется информация о значимости коэффициентов. Для каждого бита плоскости, если соответствующий коэффициент еще не является значимым и если хоть один соседний коэффициент уже является значимым, кодируется факт значимости для текущего коэффициента, т. е. фактически кодируется значения данного бита текущей кодируемой плоскости. Если кодируемый бит оказался ненулевым, то сразу после его обработки кодируется бит, соответствующий знаку этого коэффициента (кодирование знака).

Во время второго кодового прохода кодируются все биты, соответствующие значимым на данный момент коэффициентам. В отличие от предыдущего кодового прохода, когда решение о кодировании принималось на основе информации о значимости соседних коэффициентов, во время данного прохода биты кодируются в обязательном порядке.

Цель третьего кодового прохода — обработать те биты, которые не были обработаны во время первого и второго проходов. Во время третьего прохода арифметическое кодирование применяется совместно с групповым кодированием, основанным на устранении статистической избыточности представления данных.

Существенной деталью, предусмотренной стандартом, является возможность пропуска кодовых проходов, что является еще одним источником повышения эффективности за счет информационных потерь (первым, наиболее явным источником, является квантование). Данная возможность активно используется для контроля над скоростью генерации кода.

Представление информации, полученное в результате вейвлет-преобразования, очень удобно тем, что оно обеспечивает возможность получения приблизительных копий изображения без полного обратного преобразования, производимого в порядке, обратном порядку прямого преобразования. Производя ограниченное число обратных декомпозиций (объединение частотных диапазонов), полагая, что все не вовлеченные в преобразование частотные диапазоны содержат исключительно нулевые элементы, мы легко можем получить либо копию изображения в уменьшенном масштабе, либо исходное изображение, но в более низком качестве по сравнению с изображением, полученным в результате полного обратного преобразования. Учитывая тот факт, что блоки вейвлет-коэффициентов кодируются независимо друг от друга, мы получаем возможность частичного декодирования не только на уровне преобразования, но и на уровне интерпретации кода. Для получения приблизительной копии изображения достаточно декодировать лишь часть информации, а затем произвести частичное обратное преобразование. Таким образом, формат хранения изображения обеспечивает масштабируемость как по разрешению, так и по качеству.

Другим важным преимуществом нового стандарта является возможность доступа к отдельным элементам изображения без полного декодирования его представления. Обеспечивается такая возможность, во-первых, разбиением исходного изображения на непересекающиеся области, которые кодируются как отдельные изображения, а во-вторых, представлением кода отдельной области в виде частей

(слоев), каждая из которых является суммарным кодом коэффициентов, соответствующих данной области (отметим, что слои в свою очередь делятся на так называемые пакеты, содержащие код блоков коэффициентов на разных уровнях декомпозиции). Для того чтобы декодировать какую-либо область изображения достаточно определить, каким соседним областям она принадлежит и какие слои, относящиеся к этим тайлам, содержат код блоков коэффициентов, необходимых для восстановления требуемой области.

Безусловно, «удобное» представление изображения не может быть выгодным с точки зрения эффективности сжатия. Действительно, с уменьшением размера структурных элементов (областей изображения) эффективность сжатия несколько снижается. Стандарт в данном случае оставляет нам выбор: с одной стороны, мы имеем возможность получать информационные представления, позволяющие достаточно быстро извлекать и редактировать части изображения, с другой стороны, стандарт не препятствует созданию информационных представлений, эффективных по объему.

Для обеспечения помехоустойчивости информационного представления и удобства доступа к информации в стандарте JPEG 2000 предусмотрена система маркеров и маркерных сегментов. Маркеры играют роль разграничителей внутри информационного потока, маркерные же сегменты содержат в себе параметры фрагментов информации ограниченных маркерами. Данные, начинающиеся с маркера, как правило, могут быть корректно проинтерпретированы без какой-либо дополнительной информации (это, естественно, не означает возможность восстановления целого по фрагментам), что обеспечивает возможность частичного восстановления изображения, представление которого было повреждено. Введение элементов помехоустойчивости дает возможность использования стандарта во всевозможных телекоммуникационных приложениях.

Как правило, изображение включает в себя сразу несколько компонентов. Чаще всего оно состоит из трех цветовых компонентов: красного, зеленого и синего. Так как каждый компонент в отдельности представляет собой матрицу, то для того чтобы закодировать изображение целиком, необходимо закодировать не одну, а три матрицы. Такой подход, как показывает практика, приемлем, но не является самым удачным. Большей эффективности сжатия можно добиться в случае, когда кодируемые компоненты представлены в яркостно-цветовой форме.

Достижение высокого качества сжатия, безусловно, было одной из главных задач при его создании, и здесь разработчики добились явного прогресса. Стандарт JPEG 2000 превосходит по эффектив-

ности стандарт JPEG примерно в 2 раза при сжатии с потерями, являясь одним из наилучших стандартов сжатия по данному показателю на сегодняшний день, и на 5...20 % при сжатии без потерь. Основным недостатком данного стандарта является достаточно высокая вычислительная сложность составляющих его процедур, что затруднило его активное внедрения в мобильные средства, однако в сети Интернет уже можно встретить достаточно большое количество файлов, содержащих графическую информацию и обработанных в соответствии с рекомендацией данного стандарта.

6.3.2. Форматы представления неподвижного изображения

Ниже представлены векторные и растровые форматы представления сигнала неподвижного изображения.

В *векторном формате* рисунок представляется в виде комбинации простых геометрических фигур — точек, отрезков прямых и кривых, окружностей, прямоугольников и т. п. При этом для полного описания рисунка необходимо знать вид и базовые координаты каждой фигуры, например координаты двух концов отрезка, координаты центра и диаметр окружности и т. д. Этот способ кодирования идеально подходит для рисунков, которые легко представить в виде комбинации простейших фигур.

Примером изображений векторного графического формата могут служить различные чертежи, например, квартир, домов, земельных участков, дорог и т. д. В таких форматах файлов даже при большом увеличении сохраняется высокая точность, меняется только масштаб.

Растровый формат характеризуется тем, что все изображение по вертикали и горизонтали разделяется на мелкие прямоугольники — элементы изображения, или *пиксели* (pixel — picture element).

В файле, содержащем растровую графику, хранится информация о цвете каждого пикселя данного изображения. Чем меньше прямоугольники, на которые разбивается изображение, тем больше разрешение, т. е. тем более мелкие детали можно закодировать в таком графическом файле.

Информация о растровом изображении представляется в виде двумерной матрицы, элементами которой являются числа. В простейшем случае эти числа соответствуют яркостям пикселей (например, в представлениях RGB и CMYK). Кроме того, могут использоваться так называемые индексированные цвета. В этом случае числа указывают не яркость пикселей, а ячейку в таблице (палитре) цветов, которая также должна храниться в файле. Таким образом, в общем случае можно говорить о значениях пикселей, содержание

которых варьируется в зависимости от способа представления графической информации. Векторные изображения представляются в файле иначе: там нет понятия пикселей, но есть понятие объектов (линии, области заливки цветом), которые имеют свой способ описания.

Кроме описания собственно изображения в файле может находиться и другая (служебная) информация. Как конкретно записана информация в файле — это и определяется его форматом (типом).

Формат — это некоторое описание (спецификация) того, что именно, где и в каком виде должно быть представлено в файле. Для существующего формата могут создаваться как соответствующие файлы, так и программы их просмотра и редактирования. Однако некоторые графические редакторы сохраняют изображения в файлах, не реализуя всех возможностей соответствующего формата. Например, формат допускает сохранение нескольких каналов в изображении, а редактор сохраняет только один. Таким образом, следует понимать, что формат фиксирует идею представления информации, доведенную до предельно ясной спецификации, а тип файла указывает на некоторое конкретное воплощение этой идеи.

В настоящее время существует более двух десятков форматов графических файлов, например, к основным растровым форматам относятся BMP, GIF, TIF/TIFF, JPEG, PNG, RAW, WBMP, PCX, PCD, PSD, FLM, IFF, PXR, SCT/PICT, PCT, TGA, FPX, PhotoCD, MNG, ICO, FLA/SWF, а к основным векторным — WMF, EMF, CGM, EPS, WPG, AutoCAD, DXF, DWG, CDR, AI.

Создав изображение (с помощью сканера, цифрового фотоаппарата, графического редактора), необходимо сохранить его в виде файла. Для этого нужно оптимизировать параметры как самого изображения, так и файла. Главная задача при этом — найти компромисс между качеством изображения и объемом содержащего его файла.

Формат BMP. Формат BMP (*BitMaP* — битовая карта или битовый массив) представляет из себя, как правило, несжатое изображение и использует цветовую модель *RGB* для описания пикселей. Цветовая модель *RGB* наиболее часто используется при описании цветов, получаемых смешением световых лучей. Она подходит для описания цветов, отображаемых мониторами, получаемых сканерами, но не печатающими устройствами. Цвет в модели *RGB* представляется как сумма трех базовых цветов — красного (*Red*), зеленого (*Green*) и синего (*Blue*). В модели *RGB* каждый базовый цвет характеризуется яркостью (интенсивностью), которая может принимать 256 значений от 0 до 255. Таким образом, можно по-

лучить $256 \times 256 \times 256 = 16\,777\,216$ цветов. Каждому цвету можно сопоставить код, который содержит значения яркости трех составляющих. Используются десятичное и шестнадцатеричное представления кода. Десятичное представление — это тройка десятичных чисел, разделенных запятыми. Первое число соответствует яркости красной составляющей, второе — зеленой, а третье — синей. Код цвета в шестнадцатеричном представлении имеет вид $0 \times \text{XXXXXX}$. Префикс $0 \times$ указывает на шестнадцатеричное число. За префиксом следуют шесть шестнадцатеричных цифр (0, 1, 2, ..., 9, A, B, C, D, E, F). Первые две цифры — яркость красной составляющей, вторая и третья пары соответствуют яркости зеленой и синей составляющих. Если все составляющие имеют максимальную яркость (255, 255, 255 — в десятичном представлении, $0 \times \text{FFFFFF}$ — в шестнадцатеричном представлении), то получается белый цвет. Минимальная яркость (0, 0, 0 или 0×000000) соответствует черному цвету.

Базовые цвета смешиваются следующим образом:

- красный и зеленый — при максимальной яркости дают желтый цвет. Уменьшение яркости красного изменяет результирующий цвет в сторону зеленоватого, а уменьшение яркости зеленого делает цвет оранжевым;
- зеленый и синий — при максимальной яркости дают голубой. Изменяя пропорцию яркостей можно получить 65 000 оттенков голубого, от небесного до темно-синего;
- красный и синий — при максимальной яркости дают пурпурный или фиолетовый. Уменьшение яркости синего сдвигает цвет в сторону розового, а уменьшение красного — в сторону пурпурного (рис. 6.11).

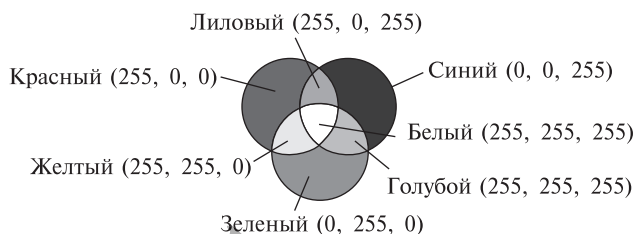


Рис. 6.11. Цветовая модель RGB

Поскольку яркость каждой из базовых составляющих цвета может принимать только 256 целочисленных значений, то каждое значение можно представить 8-разрядным двоичным числом, или одним байтом. В модели RGB информация о каждом цвете требует 3 байта (по одному на каждый базовый цвет), или 24 бита памяти для

хранения. Для представления любого из 256 оттенков серого требуется лишь 1 байт, поскольку все оттенки серого цвета образуются смешением трех составляющих одинаковой яркости.

Модель *RGB* еще называют аддитивной, поскольку по мере увеличения яркости составляющих цветов увеличивается яркость результирующего цвета.

Черно-белые файлы *BMP* называют 1-битными изображениями; изображения, которые имеют до 16 цветов, — 4-битными изображениями. Файлы формата *BMP*, имеющие 256 цветов, являются 8-битными, 65536 цветов — 16-битными и 16 миллионов цветов являются 24-битными изображениями.

В формате *BMP* есть поддержка сжатия по алгоритму *RLE*, однако теперь существуют форматы с более сильным сжатием и из-за большого объема *BMP* редко используется в сетях передачи, где для сжатия без потерь используются форматы *PNG* и более старый *GIF*. Файлы, содержащие изображения в формате *BMP*, имеют такие расширения, как *.bmp*, *.rle*, *.dib*.

Формат GIF. Формат *GIF* (*Graphics Interchange Format* — формат графического обмена) предназначен для сохранения растровых изображений с количеством цветов не более 256. В настоящее время существуют две версии формата — *GIF87a* и *GIF89a*. Имена файлов этих форматов имеют расширение *.gif*. Формат *GIF* использует алгоритм сжатия без потерь *LZW*. Метод сжатия *LZW*, разработанный в 1978 г. Лемпелом и Зивом, сжимает данные путем поиска одинаковых последовательностей (они называются фразами) во всем файле. Выявленные последовательности сохраняются в таблице, им присваиваются более короткие маркеры (ключи). Так, если в изображении имеются наборы из розового, оранжевого и зеленого пикселей, повторяющиеся 50 раз, *LZW* выявляет это, присваивает данному набору отдельное число (например, 7) и затем сохраняет эти данные 50 раз в виде числа 7. Метод *LZW* так же, как и *RLE*, лучше действует на участках однородных, свободных от шума цветов.

Формат *GIF* допускает чересстрочное хранение данных. При этом строки разбиваются на группы и меняется порядок хранения строк в файле. При загрузке изображение проявляется постепенно, в несколько проходов. Благодаря этому, имея только часть файла, можно увидеть изображение целиком, но с меньшим разрешением.

В чересстрочном *GIF* сначала записываются строки 1, 5, 9 и т. д. Таким образом, загрузив 1/4 данных, пользователь будет иметь представление о целом изображении. Вторым проходом следуют строки 3, 7, 11,..., разрешение изображения в браузере еще вдвое увеличивается. Наконец, третий проход передает все недостающие

строки (2, 4, 6...). Таким образом, задолго до окончания загрузки файла пользователь может понять, что внутри, и решить, стоит ли ждать полной загрузки изображения. Чересстрочная запись незначительно увеличивает размер файла, но это, как правило, оправдывается приобретаемым свойством.

В формате GIF89a можно сохранить не только одно, а несколько изображений, которые, например, браузер показывает друг за другом с заданными частотой и временем задержки. В результате возникает эффект анимации — движущейся картинки.

В GIF-файлах хорошо сохранять контрастные изображения без плавных цветовых переходов и шума, например логотипы, баннеры, чертежи, схемы. Другими словами, чем меньше нюансов и чем больше однородных по цвету областей в изображении, тем больше степень сжатия. Для изображений типа фотографии, когда требуется высокое качество цветопередачи, формат GIF не годится из-за ограничения количества цветов.

Формат PNG. Формат PNG — единственный из распространенных форматов, позволяющий получать полноцветные изображения с прозрачным фоном. В формате PNG использован алгоритм сжатия Deflate без потерь информации, основанный на популярном LZW-сжатии. Формат PNG не поддерживает многоканальных изображений и цветовых профилей.

Формат PNG позволяет выбирать палитру сохранения — серые полутона, 256 цветов, true color («истинные цвета»). Позволяет использовать «прозрачный» цвет, но, в отличие от GIF, таких цветов может быть до 256. В отличие от GIF, сжатие без потери качества производится и по горизонтали, и по вертикали (алгоритм собственный, параметры не настраиваемые).

PNG поддерживает три типа растровых изображений:

- полутоновое изображение (с глубиной цвета 16 бит);
- цветное индексированное изображение (палитра 8 бит для цвета глубиной 24 бит);
- полноцветное изображение (с глубиной цвета 48 бит).

Формат PNG дополнен такими возможностями, как двумерный interlacing (т.е. изображение проявляется постепенно не только по строкам, но и по столбцам) и встроенная гамма-коррекция, позволяющая сохранять изображения, яркость которых одинакова как на PC, так и на компьютерах Mac, Sun и Silicon Graphics.

Формат PNG хранит графическую информацию в сжатом виде, сжатие производится без потерь.

Он имеет следующие основные преимущества перед GIF:

- практически неограниченное количество цветов в изображении (GIF использует в лучшем случае 8-битный цвет);
- опциональная поддержка альфа-канала;
- возможность гамма-коррекции;
- двумерная чересстрочная развертка;
- возможность расширения формата пользовательскими блоками.

Формат PNG обладает более высокой степенью сжатия для файлов с большим количеством цветов, чем GIF, но разница составляет около 5...25 %, что недостаточно для абсолютного преобладания формата, так как небольшие 2–16-цветные файлы формат GIF сжимает с неменьшей эффективностью.

PNG является хорошим форматом для редактирования изображений, даже для хранения промежуточных стадий редактирования, так как восстановление и пересохранение изображения проходят без потерь в качестве. Также, в отличие, например, от TIFF, спецификация PNG не позволяет авторам реализаций выбирать, какие возможности они собираются реализовать. Поэтому любое сохраненное изображение PNG может быть прочитано в любом другом приложении, поддерживающем PNG.

Существует одна особенность GIF, которая в PNG не реализована, — поддержка множественного изображения, особенно анимации; PNG изначально был предназначен лишь для хранения одного изображения в одном файле.

Для передачи анимированных изображений был разработан расширенный формат *MNG*, опубликованный в середине 1999 г. и уже поддерживаемый в различных приложениях, однако пока так и не ставший общепринятым.

Формат TIFF. Формат TIFF (Tagged Image File Format) создан объединенными силами Aldus, Microsoft и Next специально для хранения сканированных изображений. Исключительная гибкость формата сделала его действительно универсальным. Хотя с момента его создания прошло уже много времени, TIFF до сих пор является основным форматом, используемым для хранения сканированных изображений и размещения их в издательских системах и программах иллюстрирования. Версии формата существуют на всех компьютерных платформах, что делает его исключительно удобным для переноса растровых изображений между ними. TIFF поддерживает монохромные, индексированные, полутоновые и полноцветные изображения в моделях *RGB* и *CMYK* с 8- и 16-битными каналами. Он позволяет хранить контуры, калибровочную информацию, параметры печати. Допускается использование любого количества

дополнительных альфа-каналов. Дополнительные цветовые каналы не поддерживаются.

Имеется возможность сохранять изображение в файле формата TIFF со сжатием и без сжатия. Степени сжатия зависят от особенностей самого сохраняемого изображения, а также от используемого алгоритма. Формат TIFF позволяет использовать следующие алгоритмы сжатия: PackBits (RLE), Lempel-Ziv-Welch (LZW), LZ77, ZIP, JBIG, JPEG, CCITT Group 3, CCITT Group 4. При этом JPEG является просто инкапсуляцией формата JPEG в формат TIFF. Формат TIFF позволяет хранить изображения, сжатые по стандарту JPEG, без потерь данных (JPEG-LS).

Алгоритмы CCITT Group 3 и 4 предназначены для кодирования бинарных растровых изображений. Первоначально они были разработаны для сетей факсимильной связи (поэтому иногда их называют Fax 3, Fax 4). В настоящий момент они также используются в полиграфии, системах цифровой картографии и географических информационных системах. Алгоритм Group 3 напоминает RLE, так как кодирует линейные последовательности пикселей, а Group 4 — двумерные поля пикселей.

Формат PDF. Portable Document Format (PDF) — формат электронных документов, созданный фирмой Adobe Systems с использованием ряда возможностей языка PostScript. В первую очередь предназначен для представления в электронном виде полиграфической продукции, — значительное количество современного профессионального печатного оборудования может обрабатывать PDF непосредственно. PDF с 1 июля 2008 г. является открытым стандартом ISO 32000.

Формат PDF позволяет внедрять необходимые шрифты (построчный текст), векторные и растровые изображения, включает механизм электронных подписей для защиты и проверки подлинности документов. В этом формате распространяется большое количество сопутствующей документации.

Чаще всего PDF-файл является комбинацией текста с растровой и векторной графикой, реже — текста с формами, JavaScript'ом, 3D-графикой и другими типами элементов.

Информационные объемы двух одинаково выглядящих на экране PDF-документов могут значительно отличаться. Это зависит:

- от внедрения или связывания шрифтов и мультимедиа;
- разрешения растровых изображений;
- использования встроенного механизма сжатия всего документа;
- используемых алгоритмов сжатия растровых изображений.

Для создания документа минимального объема необходимо использовать векторную графику и «безопасные» шрифты, которые можно использовать без внедрения в документ — их должны правильно отображать все программы. Любые другие шрифты, которые не были внедрены в документ и отсутствуют в системе, будут отображаться одним из имеющихся, что может стать причиной увеличения или уменьшения числа страниц, количества символов в строке, межстрочного интервала и других неприятностей, связанных с метрикой шрифта.

Формат DjVu. DjVu (от фр. *deja vu* — «уже виденное») — технология сжатия изображений с потерями, разработанная специально для хранения сканированных документов — книг, журналов, рукописей и прочее, где обилие формул, схем, рисунков и рукописных символов делает чрезвычайно трудоемким их полноценное распознавание. Такая технология является эффективным решением, если необходимо передать все нюансы оформления, например, исторических документов, где важное значение имеет не только содержание, но и цвет и фактура бумаги; дефекты пергамента: трещинки, следы от складывания; исправления, кляксы, отпечатки пальцев; следы, оставленные другими предметами.

Формат оптимизирован для передачи по сети таким образом, что страницу можно просматривать еще до завершения скачивания. DjVu-файл может содержать текстовый слой, что позволяет реализовать полнотекстовый поиск по файлу. Кроме того, DjVu-файл может содержать встроенное интерактивное оглавление и активные области — ссылки, что позволяет реализовать удобную навигацию в DjVu-книгах.

В основе формата DjVu лежат несколько технологий, разработанных в AT&T Labs:

- алгоритм отделения текста от фона на отсканированном изображении;
- вейвлетный алгоритм сжатия фона IW44;
- алгоритм сжатия черно-белых изображений JB2;
- универсальный алгоритм сжатия ZP;
- алгоритм распаковки «по запросу»;
- алгоритм «маскировки» изображений.

Для сжатия цветных изображений в DjVu применяется специальная технология, разделяющая исходное изображение на три слоя: передний план, задний план и черно-белую (однобитовую) маску. Маска сохраняется с разрешением исходного файла; именно она содержит изображение текста и прочие четкие детали. Разрешение

заднего плана, в котором остаются иллюстрации и текстура страницы, понижается для экономии места. Передний план содержит цветовую информацию о деталях, не попавших в задний план; его разрешение понижается еще сильнее. Затем задний и передний планы сжимаются с помощью вейвлет-преобразования, а маска — алгоритмом JB2.

Особенностью алгоритма JB2 является то, что он ищет на странице повторяющиеся символы и сохраняет их изображение только один раз. В многостраничных документах каждые несколько подряд идущих страниц пользуются общим «словарем» изображений.

Для сжатия большинства книг можно обойтись только двумя цветами. В этом случае используется всего один слой, что позволяет достичь высокой степени сжатия. В типичной книге с черно-белыми иллюстрациями, отсканированной с разрешением 600 dpi, средний размер страницы составляет около 15 кбайт, т.е. приблизительно в 100 раз меньше, чем исходный файл, однако при этом не стоит забывать, что в DjVu используется сжатие данных с потерями. Для особо важных документов, возможно, будет разумнее использовать более «надежные» форматы: PNG, JPEG 2000, TIFF и т.п. В общей сложности выигрыш объема в этом случае составляет 4–10 раз.

6.3.3. Принципы и особенности кодирования сигнала подвижного изображения

Качество цифрового подвижного изображения (часто используются термины «видеосигнал», «цифровой видеосигнал») определяется тремя основными факторами:

- частотой дискретизации видеопотока во временной области, определяющей количество кадров, формируемых за секунду;
- разрешением кадра (количеством пикселей в кадре);
- качеством кодирования отдельного кадра.

Существующие стандарты (алгоритмы) кодирования сигнала подвижного изображения используют различные частоты дискретизации видеопотока во временной области. Известно, что человеку для естественного восприятия подвижного изображения достаточно примерно 24 кадра в секунду. Формирование менее 10 кадров в секунду применяется для низкоскоростных визуальных сервисов (например, систем видеонаблюдения), не требующих высокого качества изображения. Некоторые устройства (например, видеотелефоны, видеодомофоны, видеорегистраторы), ориентированные на низкоскоростную передачу или запись, допускают частоты дискретизации от 10 до 25 кадров в секунду, но и в этом случае возможно прерывистое отображение движущихся объектов, создающее заметный

дискомфорт для зрителя. В телевизионных сервисах, как правило, сигнал подвижного изображения дискретизируется 25–30 раз в секунду. Еще большие частоты дискретизации (до 60 кадров в секунду), используемые в высокоскоростных аудиовизуальных сервисах (например, в телевидении высокой четкости), обеспечивают формирование в точке приема практически натурального подвижного изображения.

Разрешение кадра в настоящее время определяется множеством разрешенных вариантов. Так, достаточно часто за исходный формат принимаются обобщенный промежуточный формат *CIF* (Common Intermediate Format) и производные от него форматы, представленные в табл. 6.6.

Таблица 6.6

Форматы представления кадров

Формат	Разрешение кадра	Кол-во пикселей в кадре
SQCIF	128 пикселей × 96 строк	12288
QCIF	176 пикселей × 144 строк	25344
CIF NTSC (30 кадров/с)	352 пикселей × 240 строк	84480
PAL (25 кадров/с)	352 пикселей × 288 строк	101376
4CIF	704 пикселей × 576 строк	405504
16CIF	1408 пикселей × 1152 строк	1622016

В традиционных форматах телевидения обеспечивается соотношение сторон кадра 4:3, а в телевидении высокой четкости — 16:9 (так, формат FullHD соответствует разрешению 1920 × 1080 пикселей). В современных системах цифрового телевидения активно используются разрешения HD (1280 × 720 пикселей), 4K (4096 × 3112 пикселей), 8K (8192 × 4320 пикселей) и другие.

Качество кодирования отдельного кадра определяется процедурами квантования и кодирования, применяемыми в видеокодере, и отображается в том числе количеством бит, затрачиваемых на представление кадра.

Различные решения в отношении перечисленных факторов качества соответствуют разным требованиям к пропускной способности каналов связи. Например, рассчитаем скорость цифрового потока на выходе видеокодера, характеризующегося следующими параметрами:

- частота дискретизации во времени определяется 25 кадрами в секунду;
- используется формат 4CIF;
- используется трехцветное представление изображения в координатах *RGB* или *YCbCr* с квантованием сигналов на 256 разрешенных уровней;

- процедуры сжатия не используются:

$$V_k = 25 \times 704 \times 576 \times 3 \times 8 = 243302400 \text{ бит/с} \approx 243 \text{ Мбит/с.}$$

Полученное значение V_k свидетельствует о том, что кодирование подвижного изображения без процедур сжатия приводит к излишне высоким требованиям к пропускной способности каналов связи, что обуславливает необходимость применения процедур эффективного кодирования.

Очевидно, что методы эффективного кодирования сигнала подвижного изображения используют статистические характеристики подвижного изображения и свойства зрительного аппарата человека по его восприятию. Сжатие видеосигнала (повышение компактности цифрового представления подвижного изображения) состоит в существенном сокращении его статистической и психофизической (психовизуальной) избыточности, основные виды которой представлены ниже.

Пространственная (внутрикадровая) статистическая избыточность, состоящая в том, что значения координат (R, G, B или Y, Cr, Cb) сигнала изображения соседних пикселей внутри кадра изображения коррелированы между собой.

Энтропийная (вероятностная) статистическая избыточность, заключающаяся в существенной неодинаковости частот встречаемости разных яркостей и цветов в изображении.

Временная (межкадровая) статистическая избыточность, проявляющаяся в значительной схожести (а значит, предсказуемости) соседних кадров подвижного изображения. Действительно, фон между смежными кадрами в большинстве случаев меняется мало, а все действие определяется смещениями относительно небольших фрагментов изображения. По этой причине необходимость передачи полной информации о смежных кадрах подвижного изображения возникает только при смене сюжета, а в остальных случаях можно ограничиваться выделением и передачей разностной информации, характеризующей направления и смещения элементов изображения, появление новых объектов или исчезновение прежних. Причем такие различия могут выделяться как относительно предыдущих, так и относительно последующих кадров.

Психофизическая (психовизуальная) избыточность, заключающаяся в том, что зрительный аппарат человека имеет различную чувствительность к изменениям яркости и цветов изображения, к крупным и мелким деталям изображения, к качеству изображения в области контуров и вне их.

К основным процедурам, используемым в кодерах видеoinформации для устранения избыточности сигнала подвижного изображения, относятся следующие:

- межкадровое предсказание (с внутренней процедурой компенсации движений для подвижных элементов изображения), нацеленное на уменьшение временной избыточности;
- внутрикадровое предсказание, позволяющее уменьшить пространственную избыточность сигнала изображения;
- трансформное кодирование сигнала изображения (или ошибки его предсказания) на основе дискретного косинусного преобразования (ДКП), нацеленное на уменьшение пространственной избыточности;
- эффективное (энтропийное, экономное) кодирование, основанное на учете статистических зависимостей параметров изображения (например, кодирование вектора коэффициентов ДКП по Хаффману).

Соответствующий тип модели кодирования часто называется гибридной ДИКМ/ДКП-моделью.

В большинстве современных кодеков сигнала подвижного изображения применяются *три типа кадров, значительно отличающихся используемыми принципами кодирования и качеством*. Их определение дано в стандарте MPEG-1 (ISO/IEC 11172-2) «Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbit/s», принятом в 1993 г.:

- *I*-кадры (Intra Coded Pictures, интра-кадры) — кадры, кодируемые независимо от других на основании компромисса между высоким качеством неподвижного изображения и допустимой при этом степенью сжатия (компрессии); процедуры кодирования *I*-кадров в основном соответствуют процедурам JPEG;
- *P*-кадры (Predictive Coded Pictures) — кадры, кодируемые на основе предсказания (с учетом компенсации движения) по предыдущим *I*- или *P*-кадрам); степень сжатия *P*-кадров выше, чем степень сжатия *I*-кадров;
- *B*-кадры (Bidirectionally-predictive Coded Pictures) — кадры, кодируемые на основе двунаправленного предсказания по предыдущим и последующим *I*- или *P*-кадрам; кадры этого типа характеризуются максимальной допустимой степенью сжатия и не используются в качестве основы для процедуры предсказания.

Таким образом, кодирование *I*-кадров основано на учете пространственной, энтропийной и психофизической избыточности, кадры этого типа обладают наименьшей степенью сжатия. *P*- и *B*-кадры

предназначены для устранения межкадровой избыточности. Наличие только *I*-кадров является обязательным, их доля в общем видеопотоке в значительной степени определяет его скорость передачи. Соотношение *P*- и *B*-кадров подбирается на основе требований к сжатию сигнала изображения и допустимой сложности декодера. Очевидно, что большая доля *B*-кадров в видеопотоке повышает как степень сжатия сигнала, так и вычислительную сложность декодера.

Естественно, что любая видеопоследовательность начинается с *I*-кадра. Очередность расположения указанных типов кадров в выходной видеопоследовательности многовариантна и зависит от нескольких факторов, в том числе — от реализуемого аудиовизуального сервиса. Пример формирования последовательности с используемыми видами предсказания показан на рис. 6.12.

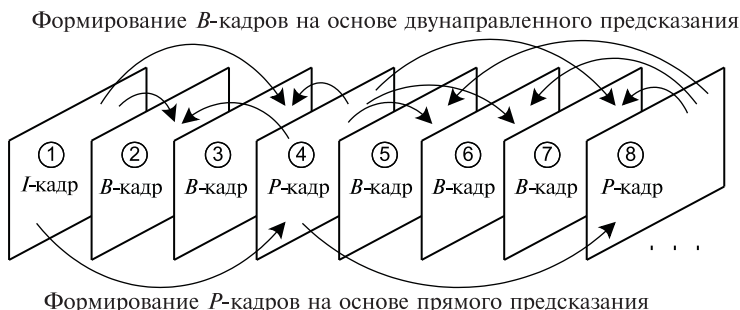


Рис. 6.12. Пример формирования последовательности *I*-, *P*- и *B*-кадров

Очередность формирования кадров в кодере (1, 4, 2, 3, 8, 5, 6, 7) требует такой же последовательности их декодирования, что обуславливает необходимость передачи декодеру информации о типе каждого кадра подвижного изображения.

Для эффективного использования процедуры компенсации движения необходимо предварительное выделение в кадре движущихся областей (элементов). Как правило, с этой целью используются следующие решения:

- структурное разделение кадров всех типов на отдельные квадратные макроблоки, объединяющие в себя несколько блоков 8×8 пикселей (свойственных алгоритму JPEG); формирование типового макроблока, определенного Рекомендацией H.262/MPEG-2 и состоящего из четырех блоков (256 пикселей), показано на рис. 6.13.
- классификация макроблоков на подвижные и неподвижные на основе сравнения соседних *I*-кадров;

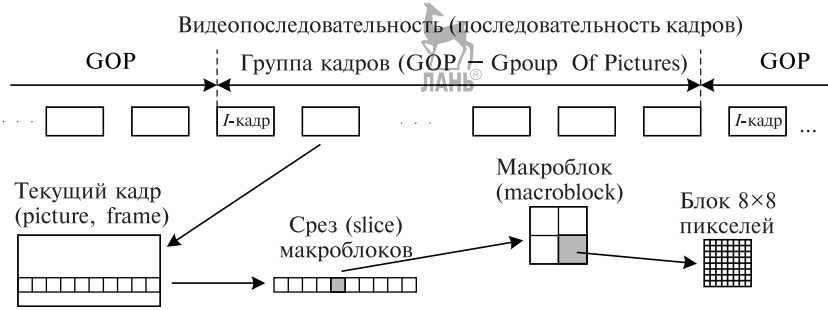


Рис. 6.13. Принцип формирования макроблока

- сохранение неподвижных макроблоков *I*-кадров в памяти для дальнейшего использования в процедурах обработки (например, экстраполяции неподвижных элементов в процедуре декодирования);
- использование подвижных макроблоков для расчета векторов движения с целью формирования *P*- и *B*-кадров.

Вариативность представления макроблоков в различных стандартах кодирования сигнала подвижного изображения обусловлена использованием различных соотношений частот дискретизации координат Y , Cb и Cr (различных форматов представления цветоразностных сигналов). Наиболее популярными форматами являются форматы 4:4:4, 4:2:2 и 4:2:0. Следует иметь в виду, что прямой трактовки цифр, использованных в обозначениях форматов (например, 4:2:0), не существует. В табл. 6.7 показаны некоторые характеристики указанных форматов.

Таблица 6.7

Некоторые характеристики форматов представления цветоразностных сигналов

Характеристика	Обозначение формата		
	4:4:4	4:2:2	4:2:0
Отношение частот дискретизации сигналов Cb и Cr к частоте дискретизации сигнала Y	1:1	1:2	1:4
Средние затраты на представление одного пикселя изображения (Y , Cb , Cr) при 256-уровневом (8-битном) квантовании, бит/пиксель	24	16	12
Затраты на представление макроблока 16×16 пикселей изображения (Y , Cb , Cr), бит	6144	4096	3072

На рис. 6.14 представлена структура информационного отображения макроблока для форматов 4:4:4, 4:2:2 и 4:2:0.

Например, в формате 4:2:0 макроблок содержит 256 отсчетов сигнала Y , 64 отсчета сигнала Cb и 64 отсчета сигнала Cr . Сле-

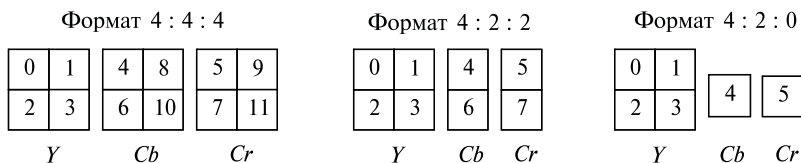


Рис. 6.14. Структура информационного отображения макроблока при различных форматах представления цветоразностных сигналов

довательно, в этом случае до применения процедуры эффективного кодирования при стандартном 256-уровневом квантовании цифровое представление макроблока содержит $(256 + 64 + 64) \times 8 \text{ бит} = 3072 \text{ бита}$. Общие информационные затраты на кодирование кадра в этом случае будут зависеть от количества макроблоков в кадре (от разрешения кадра). Например, для формата 4CIF (704×576 пикселей, 1584 макроблока) они составят 4 866 048 бит.

К настоящему времени разработано большое число стандартов компактного цифрового представления сигнала подвижного изображения, наиболее популярные из них подготовлены специалистами двух организаций:

- группы экспертов по кодированию подвижного изображения (MPEG — Moving Picture Coding Experts Group), образованной в 1988 г. Международной организацией по стандартизации (International Organization for Standardization — ISO) и Международной электротехнической комиссией (International Electrotechnical Commission — IEC); соответствующие алгоритмы (стандарты) составляют серию MPEG;
- сектора стандартизации в области телекоммуникаций Международного союза электросвязи (Telecommunication Standardization Sector of International Telecommunications Union — ITU-T); соответствующие стандарты (рекомендации) составляют серию H.26x. В 1997 г. для активизации работ по разработке стандартов кодирования сигналов подвижного изображения в ITU-T была создана группа экспертов по видеокодированию (VCEG — Video Coding Experts Group). К настоящему времени серия H.26x содержит стандарты H.261, H.262, H.263, H.264 и H.265, их полный текст размещен на сайте ITU-T www.itu.int.

В 2001 г. для улучшения координации работ по разработке новых стандартов видеокодирования MPEG и VCEG создали объединенную видеокоманду JVT (Joint Video Team).

На рис. 6.15 показана хронология разработки стандартов указанных серий (показаны даты принятия первых версий документов), а в дальнейшем учебном материале рассмотрены особенности коди-

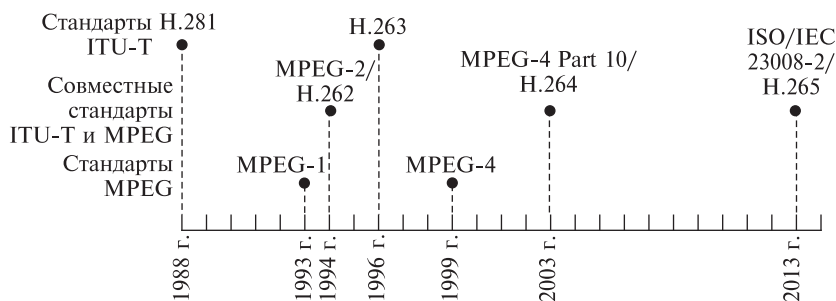


Рис. 6.15. Хронология разработки стандартов кодирования сигнала подвижного изображения

рования сигнала подвижного изображения в некоторых из представленных стандартов.

Необходимо отметить, что существуют различные способы (форматы) записи и хранения видеофайлов, сопровождающих их аудиофайлов, сформированных на основе стандартов серий MPEG и H.26x, и необходимой служебной информации. На практике применение таких способов приводит к формированию различных *видео-контейнеров*, определяющих структуру мультимедийных данных. К числу наиболее распространенных видеоконтейнеров относятся AVI, MKV, MOV, MPEG, 3GP.

6.3.4. Стандарты кодирования сигнала подвижного изображения серий H.26x и MPEG

Различные стандарты (рекомендации) кодирования сигнала подвижного изображения ориентированы на достижение компромисса между требуемым уровнем качества цифрового подвижного изображения и допустимой скоростью передачи цифрового потока (скоростью кодирования). Очевидно, что в основе совершенствования процедур кодирования, как правило, лежит повышение качества анализа сигнала, обуславливающее повышение вычислительной сложности операций и, следовательно, повышение требований к производительности процессоров, реализующих стандарты.

Стандарты серии MPEG и H.26x определяют требования к процедурам кодирования как сигнала подвижного изображения, так и аудиосигнала, связанного с изображением. Кроме того, некоторые из стандартов описывают применяемые процедуры помехоустойчивого кодирования. В материале, представленном ниже, рассмотрены лишь вопросы кодирования сигнала подвижного изображения. Следует иметь в виду, что стандарты H.261, MPEG-1, MPEG-2/H.262 и H.263, хронологически следуя друг за другом, основаны на единой

базовой модели кодера, используют общую стратегию кодирования и отображают процесс ее последовательного совершенствования. По этой причине стандарты H.261 и MPEG-1 будут представлены достаточно кратко, основное внимание будет уделено стандарту MPEG-2/H.262, а стандарт H.263 будет охарактеризован лишь перечнем усовершенствований относительно MPEG-2/H.262.

Рекомендация H.261. Рекомендация H.261 «Video codec for audiovisual services at $p \times 64$ kbit/s» принята в ноябре 1988 г. и определяет методы кодирования и декодирования сигнала подвижного изображения для применения в аудиовизуальных сервисах, использующих цифровой поток со скоростью передачи, кратной 64 кбит/с (от 64 до 30×64 кбит/с). В настоящее время действует ее редакция от марта 1993 г.

Структурная схема кодера, предложенная в H.261, легла в основу более поздних стандартов и будет рассмотрена при описании кодера MPEG-2.

Входное изображение, поступающее на вход кодера источника, представлено следующим образом:

- частота кадров не более 30 кадров в секунду;
- используются форматы CIF или QCIF;
- используется система координат (Y, Cb, Cr) ; кодирование значений координат сигнала производится в соответствии с Рекомендацией Сектора радиосвязи Международного союза радиосвязи ITU-R BT. 601-7 «Студийные параметры кодирования цифрового телевидения для стандартного 4:3 и широкоэкрannого 16:9 форматов», принятой в марте 2011 г.

При кодировании используется формат представления цветоразностных сигналов 4:2:0. Обеспечиваемая точность компенсации движений составляет один пиксель.

MPEG-1. Стандарт MPEG-1 (ISO/IEC 11172-2) «Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbit/s» был принят в 1993 году. Предложенный в нем алгоритм кодирования сигнала подвижного изображения предназначен для записи цифрового видеосигнала и ассоциированного с ним аудиосигнала на компакт-диски CD. Основные процедуры кодирования и структура кодера в основном соответствуют стандарту H.261, скорость кодирования сигнала подвижного изображения составляет около 1,5 Мбит/с.

Максимально допустимое разрешение кадра составляет $768 \text{ пикселей} \times 576 \text{ строк}$, частота кадров не более 30 кадров в секунду. При кодировании используется формат представления цветоразностных сигналов 4:2:0.

Размер макроблока, используемого в процедуре компенсации движений, составляет 16×16 пикселей.

Максимально возможная скорость выходного цифрового потока составляет 1,86 Мбит/с.

В настоящее время использование стандарта продолжается, однако по популярности и качеству цифрового сигнала подвижного изображения он уступает более поздним стандартам серии MPEG.

MPEG-2/Н.262. Стандарт MPEG-2 (ISO/IEC 13818-2)/Н.262 «Generic coding of moving pictures and associated audio information» был принят в 1994 г. и определил метод универсального кодирования сигнала подвижного изображения и ассоциированного аудио для широкого диапазона аудиовизуальных сервисов (запись видеосигнала на диски *DVD*, использование в разновидностях цифрового телевидения и проч.). В настоящее время после некоторых усовершенствований данный стандарт действует в редакции от февраля 2000 г.

Стандарт предусматривает поддержку нескольких стандартизированных профилей и уровней функционирования, различающихся типом и разрешением кадра, частотой следования кадров, форматом передачи цветоразностных сигналов, скоростью выходного цифрового видеопотока (табл. 6.8, 6.9). Комбинирование профилей и уровней и возможность изменения значений их параметров обуславливают значительную гибкость стандарта, обеспечивающую возможность его применения в большом диапазоне аудиовизуальных сервисов. Различные частоты дискретизации видеопотока во времени позволяют формировать от 23,98 до 60 кадров в секунду.

Таблица 6.8

Профили стандарта MPEG-2

Профиль	Типы кадров	Формат YCbCr	Особенности
Simple Profile (SP)	I, P	4:2:0	Не поддерживается чресстрочная развертка
Main Profile (MP)	I, P, B	4:2:0	Учитывается ОСШ Поддержка трех уровней качества декодирования Обеспечение максимально возможного качества декодирования
422 Profile (422P)	I, P, B	4:2:2	
SNR Profile (SNRP)	I, P, B	4:2:0	
Spatial Profile (SpP)	I, P, B	4:2:0	
High Profile (HP)	I, P, B	4:2:2	

Обобщенная структурная схема кодера сигнала подвижного изображения стандарта MPEG-2 представлена на рис. 6.16.

Процесс кодирования изображения можно условно представить следующими основными этапами и процедурами:

Таблица 6.9

Уровни стандарта MPEG-2

Уровень	Разрешение кадра, пикселей	Временная дискретизация, кадров / с	Скорость выходного цифрового видеопотока, Мбит / с
Low Level (LL)	352×288	30	4
Main Level (ML)	720×576	30	15
High 1440 (H-14)	1440×1152	30	60
High Level (HL)	1920×1152	30	80

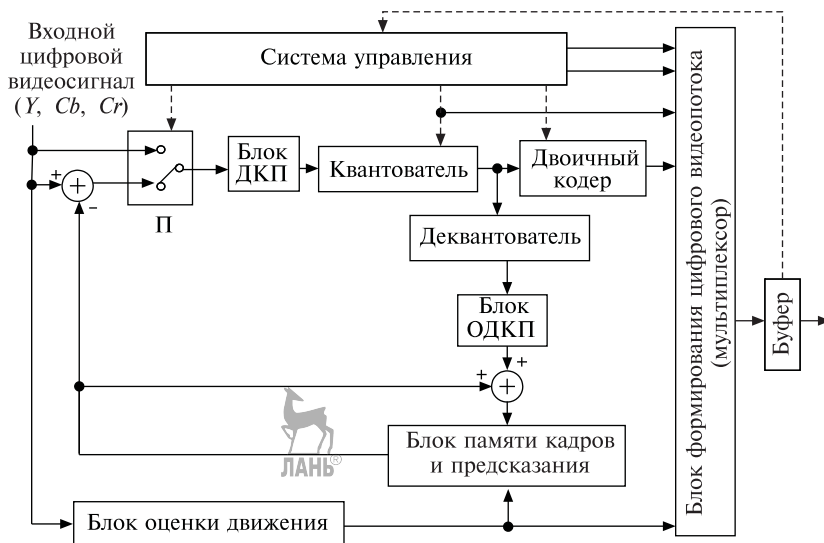


Рис. 6.16. Обобщенная структурная схема кодера стандарта MPEG-2

Определение статических и подвижных элементов изображения, формирование векторов движения. На основе сравнения текущего и предыдущих I -кадров кодер определяет статические элементы изображения и элементы, подверженные перемещению. Для движущихся элементов (макроблоков) рассчитываются векторы движения, отображающие направление и скорость перемещения. В последующем статические элементы изображения и рассчитанные векторы движения используются в процедурах кодирования (декодирования) P - и B -кадров.

Кодирование I-кадров. Режим внутрикадрового кодирования, соответствующий кодированию I-кадров, формируется установкой переключателя П в верхнее положение. В данном режиме макроблоки входного кадра цифрового изображения в координатах (Y, Cb, Cr) последовательно подвергаются процедурам дискретного

косинусного преобразования (ДКП), неравномерного квантования и двоичного кодирования, что в целом соответствует алгоритму JPEG. Кодовая комбинация, отображающая I -кадр, с выхода двоичного кодера поступает на блок формирования цифрового видеопотока. Система управления кодером обеспечивает формирование I -кадров примерно каждые 0,5 с. На выходе блока обратного ДКП формируется квантованный кадр исходного изображения в координатах (Y , Cb , Cr), отличающийся от оригинального на ошибку квантования и соответствующий кадру, который будет сформирован в декодере MPEG-2. Данный кадр поступает в блок памяти и предсказания для использования в процедурах кодирования P - и B -кадров.

При кодировании P - и B -кадров переключатель Π устанавливается в нижнее положение, что означает, что кодированию подвергается не исходный кадр изображения, а разность между ним и его предсказанным значением.

Кодирование P -кадров. В процедуре предсказания используются несколько предыдущих квантованных I - и/или P -кадров и актуальные векторы движения, позволяющие корректировать предсказание подвижных элементов изображения. Таким образом, динамический диапазон сигнала, поступающего в тракт ДКП, квантования и двоичного кодирования, существенно меньше, чем для I -кадров, что обуславливает большую степень сжатия.

Естественно, что для декодирования P -кадров необходимо передать декодеру не только кодовую комбинацию, формируемую на выходе двоичного кодера, но и информацию об используемых параметрах предсказания и векторе движения.

Кодирование B -кадров. Применение в процедуре кодирования B -кадров двунаправленного предсказания позволяет обеспечить еще большую степень компрессии. В общем случае B -кадр содержит лишь информацию об изменении состояния подвижных элементов изображения.

Таким образом, с учетом разного подхода к кодированию I -, P - и B -кадров, на выходе двоичного кодера формируется цифровой поток, имеющий переменную скорость.

Формирование цифрового видеопотока. Очевидно, что многовариантность кодирования сигнала в кодере MPEG-2 требует передачи декодеру не только цифрового отображения кадров, формируемого двоичным кодером, но и иной информации, необходимой декодеру для верного декодирования сигнала подвижного изображения. С этой целью в блоке формирования цифрового видеопотока производится временное объединение (мультиплексирование) цифрового потока с выхода двоичного кодера и служебной информации,

поступающей с различных структурных элементов кодера. Таким образом, в видеопотоке на выходе мультимплексора структурно объединены следующие данные:

- актуальные параметры используемого профиля;
- актуальные параметры используемого уровня;
- таблица квантования;
- цифровое отображение последовательности кадров с выхода двоичного кодера;
- векторы движения;
- служебная информации о параметрах GOP;
- служебная информация о кадрах (типы кадров);
- служебная информация о срезах макроблоков;
- служебная информация о макроблоках;
- иная служебная информация.

Выравнивание скорости цифрового видеопотока. На выходе мультимплексора, как и на выходе двоичного кодера, формируется цифровой поток, имеющий переменную скорость. Для формирования цифрового потока с постоянной скоростью передачи (выравнивания скорости передачи) в кодере MPEG-2 используется буфер, представляющий собой запоминающее устройство значительного объема. Запись данных в буфер производится по мере их поступления с выхода мультимплексора. Считывание информации из буфера начинается после некоторого его заполнения и происходит в режиме *FIFO* (First Input First Output, «первым вошел — первым вышел») на постоянной скорости. Для устойчивой работы буфера в таком режиме требуется поддерживать степень его заполнения в определенном диапазоне значений. С этой целью производится оценка заполнения буфера, значение которой через систему управления регулирует параметры квантователя, увеличивая или уменьшая точность квантования коэффициентов ДКП. Такая процедура выравнивания скорости видеопотока одновременно обеспечивает желательное соотношение между воспринимаемой глазом мерой качества изображения и степенью его активности:

- при высокой активности изображения зрению достаточно образов более низкого качества, что и происходит — скорость цифрового потока на выходе кодера возрастает, степень наполнения буфера возрастает, система управления дает команду квантователю на понижение точности квантования, в результате качество изображения уменьшается;
- при низкой активности изображения зрение более требовательно к его качеству, что и обеспечивается — скорость цифрового

потока на выходе кодера уменьшается, степень заполнения буфера также уменьшается, при этом система управления дает команду квантователю на повышение точности квантования, в результате качество изображения возрастает.

Рекомендация Н.263. Рекомендация Н.263 «Video coding for low bit rate communication» принята в марте 1996 г. и определяет методы кодирования и декодирования подвижного изображения для применения в аудиовизуальных сервисах, использующих низкие скорости передачи (в настоящее время действует редакция от января 2005 года). Рекомендация расширяет возможности Н.261, что приводит к усложнению процедур обработки сигнала. Рекомендация не содержит ограничений по скорости формируемого цифрового видеопотока, данные ограничения накладываются используемым терминальным оборудованием или сетью передачи данных.

Базовая конфигурация кодера основана на Рекомендации Н.261. В качестве форматов входного изображения используются SQCIF, QCIF, CIF, 4CIF, 16CIF, а также допускается применение иных пользовательских форматов. Точность компенсации движений составляет половину пикселя.

Основные усовершенствования кодера Н.263 по отношению к кодеку Н.261 заключаются в ряде дополнительных опций, которые могут использоваться как совместно, так и по отдельности. Некоторые из этих опций представлены ниже:

- возможность постоянной работы в многоточечном режиме с использованием видеомультиплексирования (возможность передачи в одном видеоканале до четырех независимых бинарных «видеосубканалов»);
- режим с неограниченным вектором движения (обеспечение лучшей компенсации движений на границах кадра в случае движения камеры и при больших форматах изображений);
- режим арифметического кодирования на основе синтаксиса, обеспечивающий постоянную скорость кодирования видеопотока;
- режим улучшенного предсказания, использующий компенсацию движений перекрывающихся блоков и обеспечивающий улучшение субъективного восприятия подвижного изображения;
- режимы использования *PB*- и улучшенных *PB*-кадров, что обеспечивает возможность взаимодействия с видеокодеками, использующими такие кадры;
- режим деблокирующей фильтрации, обеспечивающий уменьшение вероятности появления артефактов на границах блоков 8×8 пикселей *I*- и *P*-кадров;

- режим улучшенного кодирования I -кадров;
- режим использования дополнительной вспомогательной информации, расширяющий возможности взаимодействия кодеков H.263 с иными кодеками и внешними устройствами;
- режим выбора опорного изображения, улучшающий качество видеосвязи по каналу с ошибками;
- режим масштабирования, обеспечивающий возможность адаптации алгоритма кодирования в сетях с переменной пропускной способностью;
- режим пространственной передискретизации опорного изображения;
- режим модифицированного квантования, позволяющий повысить возможности по управлению скоростью передачи цифрового видеопотока за счет изменения шага квантования.

Совершенствование процедур кодирования сигнала подвижного изображения традиционно направлено на задачи повышения качества изображения при сохранении прежней скорости кодирования или снижения требуемой скорости кодирования при прежнем качестве изображения. Отражением этой тенденции является процесс стандартизации более совершенных алгоритмов сжатия видеосигнала. Существенно, что каждый последующий стандарт видеокодирования в различных режимах своего функционирования способен обеспечивать решение обеих из указанных выше задач.

MPEG-4. Стандарт MPEG-4 (ISO/IEC 14496) «Coding of audio-visual objects» был принят в 1999 г. и определил новый подход к кодированию сигнала подвижного изображения, в основе которого лежит предварительное разделение подвижного изображения и ассоциированного с ним аудио на отдельные элементы:

- несколько классов аудиовизуальных естественных и синтетических объектов (аудиообъект, видеообъект, аудиовизуальный объект, графический плоский объект, графический трехмерный объект, текстовый объект и др.);
- сцена как совокупность пространственно-временных отношений между аудиовизуальными объектами.

Раздельное независимое эффективное кодирование указанных элементов с возможностью масштабирования качества их представления позволяет достигать значительной вариативности в степени сжатия сигнала и используемой скорости передачи (от 5 кбит/с до 50 Мбит/с), что, в свою очередь, определяет широкую область применения стандарта: мультимедийные компьютерные сервисы, телевидение и киноиндустрия, телекоммуникационные сервисы.

Достоинством стандарта является возможность обеспечения интерактивности со стороны пользователя, заключающейся в манипуляциях с описанием сцены и с аудиовизуальными объектами на ней (например, приближение и удаление объектов, выбор камеры и направления просмотра, повторный просмотр отдельных эпизодов).

Стандарт определяет набор профилей, определяющих используемые процедуры кодирования сигнала подвижного изображения (визуальные профили, графические профили, профили графики сцены). Например, различные визуальные профили учитывают особенности реального, синтетического и смешанного видео.

Существенно, что одна из версий MPEG-4 была подготовлена совместно специалистами MPEG и ITU и представлена в документах обеих организаций: MPEG-4 Part 10 (ISO/IEC 14496-10) и H.264.

Рекомендация H.264/MPEG-4 Part 10. Рекомендация H.264 «Advanced video coding for generic audiovisual services» была принята в мае 2003 г., в настоящее время действует ее редакция от февраля 2014 г. Стандарт определяет усовершенствованный метод кодирования сигнала подвижного изображения для применения в широком классе аудиовизуальных сервисов (телевизионное вещание, видеоконференцсвязь, видеотелефония, хранение видеосигнала на оптических дисках и др.). Данная рекомендация является полным аналогом версии MPEG-4 Part 10 (ISO/IEC 14496-10). Кодеки, реализующие рекомендацию H.264, часто называются AVC-кодеками.

Рекомендация поддерживает различные форматы представления кадров и форматы представления цветоразностных сигналов (4:2:0, 4:2:2, 4:4:4) и, соответственно, широкий диапазон возможных скоростей выходного цифрового потока.

Структурная схема кодера H.264 (рис. 6.17) в основном соответствует структурным схемам предыдущих стандартов, единственным принципиальным отличием является наличие деблокирующего фильтра. Стандарт предусматривает высокую вариативность процедур обработки сигнала, что обуславливает многорежимность кодирования. При этом следует иметь в виду, что вычислительные процедуры, используемые при реализации кодирования, являются более совершенными и обеспечивают достижение лучших характеристик, чем в предыдущих стандартах.

Кодер традиционно использует режимы внутрикадрового и межкадрового предсказаний (ВКП и МКП соответственно), включая процедуры компенсации движения при МКП. Кодруемый кадр сигнала подвижного изображения предварительно разделяется на макроблоки размером 16×16 пикселей, при этом отдельные внутренние процедуры кодирования предполагают разделение макроблоков

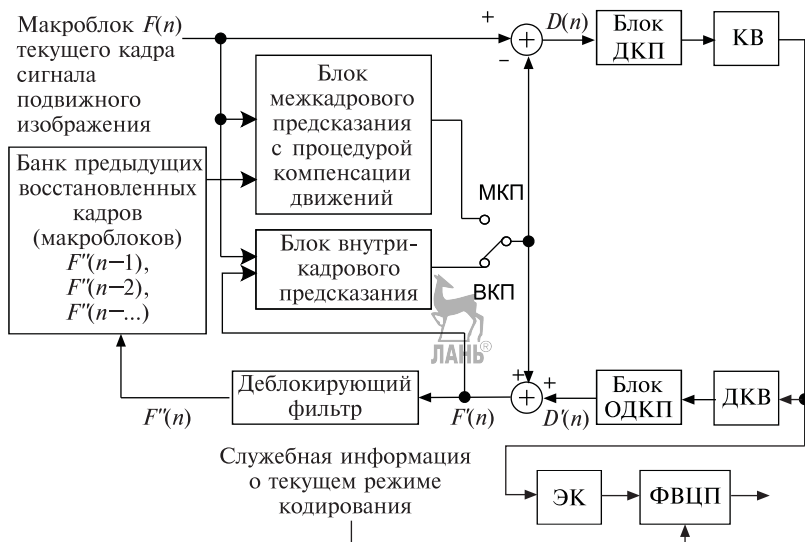


Рис. 6.17. Структурная схема кодера H.264/MPEG-4 Part 10

на блоки меньших размеров (до 4×4 пикселя). H.264 значительно превосходит предыдущие стандарты по глубине МКП: количество предыдущих кадров, которые могут использоваться в процедуре предсказания, может достигать 32-х.

На схеме показаны текущий макроблок $F(n)$ сигнала подвижного изображения, восстановленный макроблок $F'(n)$ на входе деблокирующего фильтра и восстановленный макроблок $F''(n)$ на выходе деблокирующего фильтра. Назначением деблокирующего фильтра является сглаживание искажений, возникающих на границах макроблоков вследствие их отдельной обработки.

Каждый макроблок может кодироваться в режимах ВКП или МКП с учетом следующего:

- в режиме ВКП используются восстановленные макроблоки $F'(n)$ текущего кадра, сформированные на основе квантованной ошибки предсказания $D'(n)$ текущего кадра; для формирования $D'(n)$ применяются процедуры деквантования и обратного дискретного косинусного преобразования;
- в режиме МКП используются восстановленные макроблоки $F''(n)$ предыдущих кадров.

В кодере реализуется усовершенствованный режим компенсации движений, применяющий разделение макроблоков на прямоугольные субблоки различной размерности с расчетом отдельных векторов движения для каждого субблока. Очевидно, что повыше-

ние количества субблоков (векторов движения) влечет за собой как точность МКП макроблока (уменьшение мощности $D(n)$), так и увеличение объема служебной информации о текущем режиме кодирования, а в целом оказывает существенное влияние на эффективность кодирования. Кодер Н.264 обеспечивает точность компенсации движений в одну четверть пикселя, что особенно важно для корректного отображения медленно движущихся объектов.

Выходной цифровой поток (ВЦП) формируется в формирователе ВЦП (ФВЦП) путем объединения (мультиплексирования) следующих двоичных цифровых последовательностей:

- представляющей текущий кадр (макроблок), которая формируется последовательными процедурами спектрального преобразования (СП) ошибки предсказания $D(n)$ текущего макроблока, квантования (КВ) коэффициентов СП, сканирования (С) и энтропийного кодирования (ЭК) квантованных коэффициентов СП; при этом информационное содержание $D(n)$ определяется текущим режимом кодирования;
- содержащей служебную информацию о текущем режиме кодирования (вариант предсказания, таблица квантования, способ формирования и параметры векторов движения и т. д.).

Рекомендация Н.264, в отличие от более ранних стандартов кодирования сигнала подвижного изображения, использовавших дискретное косинусное преобразование (ДКП), предусматривает три варианта СП для различных вариантов $D(n)$. Например, для формата представления цветоразностных сигналов 4:2:0 СП осуществляется следующим образом:

- обрабатываемый макроблок 16×16 пикселей разделяется на субблоки по 4×4 пикселя, к каждому из которых применяется двумерное преобразование, подобное двумерному ДКП и часто называемое приближенным ДКП; в результате преобразования формируются матрицы 4×4 коэффициентов двумерного приближенного ДКП для сигнала Y (номера 0–15 на рис. 6.18), сигнала Cb (номера 18–21 на рис. 6.18) и сигнала Cr (номера 22–25 на рис. 6.18);
- из постоянных составляющих матриц с номерами 0–15 формируется матрица размером 4×4 (номер —1 на рис. 6.18), к которой дополнительно применяется двумерное преобразование Адамара; данная процедура используется только для макроблоков, кодируемых в режиме ВКП;
- из постоянных составляющих матриц с номерами 18–21 и 22–25 формируются две матрицы размером 2×2 (номера 16 и 17

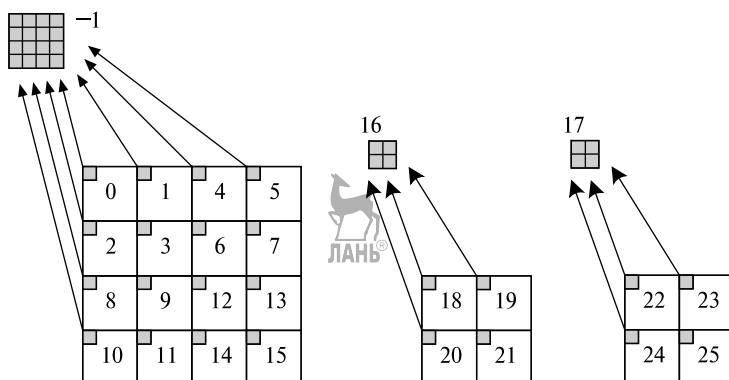


Рис. 6.18. Структурирование макроблока для формата 4:2:0 при выполнении спектрального преобразования

на рис. 6.18), к каждой из которых дополнительно применяется двумерное преобразование Адамара.

Для примера, показанного на рис. 6.18, очередность дальнейшей обработки коэффициентов СП (их общее количество равно 384) соответствует номерам матриц (последовательно от -1 до 25).

Режим сканирования коэффициентов преобразования (формирование вектора коэффициентов преобразования) многовариантен и определяется текущими вариантами предшествующих процедур. Стандарт предусматривает применение различных видов энтропийного кодирования в зависимости от кодируемых параметров и режима кодирования, в том числе контекстного адаптивного арифметического кодирования и экспоненциального кодирования Голomba.

Служебная информация о режиме кодирования может кодироваться как кодами с фиксированной размерностью кодового слова, так и кодами переменной длины.

В настоящее время Рекомендация H.264 является доминирующей в области кодирования сигнала подвижного изображения в подавляющем большинстве аудиовизуальных сервисов. Видеокодеки, реализующие рекомендацию, являются более ресурсозатратными с точки зрения вычислительной сложности кодирования и требуемой емкости памяти, но обеспечивают значительное превосходство над кодеками более ранних стандартов с точки зрения достигаемого компромисса между качеством цифрового подвижного изображения и скоростью кодирования. Существенно, что стандарт предусматривает возможность разработчикам аппаратных и программных видеокодеков использовать авторские алгоритмы реализации отдельных процедур кодирования при сохранении основных параметров кодирования и декодирования сигнала подвижного изображения.

Рекомендация Н.265. Рекомендация Н.265 «High efficiency video coding» была принята в апреле 2013 г., в настоящее время действует ее редакция от октября 2014 г. Кодеки Н.265 часто обозначаются аббревиатурой, составленной на основе слов, входящих в название рекомендации — HEVC-кодеки. Стандарт определяет усовершенствованный метод кодирования сигнала подвижного изображения, имеет значительный диапазон параметров кодирования (качества изображения, скорости кодирования) и ориентирован на применение в широком классе аудиовизуальных сервисов (телевизионное вещание, видеоконференц-связь, видеотелефония, хранение видеосигнала на оптических дисках и др.). Данная рекомендация является результатом совместной работы MPEG и VCEG и представлена в текстах Н.265 и ISO/IEC 23008-2.

Базовая модель кодирования по Рекомендации Н.265 соответствует ранее разработанным стандартам. Рекомендация обеспечивает поддержку форматов представления кадров вплоть до 8K (8192×4320 пикселей) — самого высокого из современных телевизионных форматов и форматов представления цветоразностных сигналов 4:2:0, 4:2:2 и 4:4:4.

К числу основных особенностей Н.265 следует отнести:

- возможность увеличения размеров макроблоков с 16×16 до 64×64 пикселей;
- возможность параллельного декодирования различных областей текущего кадра;
- произвольный доступ к изображениям в процессе декодирования (отсутствие необходимости наличия информации о предыдущих кадрах при декодировании произвольно выбранного кадра).

Эти и иные новшества Н.265 обуславливают повышение вычислительной сложности процедур кодирования и декодирования, что влечет за собой повышение требований к производительности процессоров, используемых при программной реализации кодеков. Однако указанные инновационные решения обеспечивают большую эффективность кодирования (достижение лучших соотношений скорости кодирования и качества цифрового подвижного изображения) по отношению к более ранним стандартам. Например, кодек позволяет передавать высококачественный сигнал подвижного изображения с разрешением кадра 4K на скоростях 20–30 Мбит/с. На рис. 6.19 приведены показатели эффективности кодирования сигнала подвижного изображения для кодеков MPEG-2, Н.264 и Н.265 при формировании 24 кадров в секунду и разрешении кадра 1920×1080 пикселей.



Рис. 6.19. Эффективность кодирования сигнала подвижного изображения кодеками MPEG-2, H.264 и H.265

В настоящее время организация мультимедийной связи в режиме реального времени (в том числе видеосвязи, видеоконференц-связи) основывается на следующих основных стандартах МСЭ-Т:

- Стандарт H.323 был принят в 1996 г. и имел название «Visual telephone systems and equipment for local area networks which provide a non-guaranteed quality of service», в настоящее время действует редакция стандарта от 2009 г. под названием «Packet-based multimedia communications systems». Стандарт определяет условия организации мультимедийных услуг связи в режиме реального времени на сетях с пакетной коммутацией, в том числе в условиях негарантированного качества обслуживания.
- Стандарт H.324 «Terminal for low bit-rate multimedia communication» принят в 2009 г. и определяет терминал для низкоскоростной мультимедийной связи.

В табл. 6.10 показано применение рассмотренных выше стандартов кодирования речевого сигнала и сигнала подвижного изображения в системах мультимедийной связи, при этом жирным шрифтом показаны основные стандарты. Использование конкретного набора дополнительных кодеков часто регламентируется стандартами H.241 «Extended video procedures and control signals for H.300-series terminals» и H.245 «Control protocol for multimedia communication».

Таблица 6.10

Применение стандартов кодирования речевого сигнала и сигнала подвижного изображения в системах видеоконференц-связи

Стандарты построения систем видеоконференц-связи	H.323	H.324
Стандарты кодирования речевого сигнала	G.711, G.722, G.723.1, G.728, G.729	G.723.1, G.711, G.722.1, G.728, G.729
Стандарты кодирования сигнала подвижного изображения	H.261, H.263, H.264	H.263, H.261, H.264

7 Цифровая модуляция и демодуляция

В настоящее время большая часть информации, передаваемой по разнообразным каналам связи, существует в цифровом виде. Это означает, что передаче подлежит последовательность целых чисел n_0, n_1, \dots , которые могут принимать значения из некоторого фиксированного конечного множества. Эти числа, называемые символами, поступают от источника информации с интервалом T_c , представляя собой дискретный сигнал, а частота поступления данных элементов, соответствующая в простейшем случае частоте дискретизации, называется символьной скоростью (symbol rate) $f_s = 1/T_c$.

Указанные символы принимают значения из конечного множества, что означает квантование сигнала, следовательно, его можно назвать цифровым сигналом.

Часто используемым на практике вариантом является двоичная (binary) последовательность символов, когда каждое из чисел может принимать одно из двух значений: 0 или 1.

При передаче дискретной последовательности символов для согласования с характеристиками канала связи с каждым из возможных значений символа сопоставляется некоторый набор параметров несущего колебания. Эти параметры поддерживаются постоянными в течение интервала T_c , т. е. до прихода следующего символа.

Данный подход позволяет ввести определение цифровой модуляции (ЦМ), которую в самом общем смысле можно определить как изменение одного или нескольких параметров несущего (модулируемого) сигнала по закону первичного цифрового (модулирующего) сигнала, представляющего информационную последовательность. При ЦМ преобразует исходный цифровой сигнал (цифровые символы, содержащие информацию, как правило, с выхода кодера источника или канала) к виду, пригодному для передачи по каналу связи (совместимому с характеристиками канала).

Как определено выше, основная цель ЦМ состоит в отображении информационной двоичной последовательности в сигналы, со-

вместимые с каналом связи, что определяет необходимость введения терминов: символ, алфавит, позиционность. При ЦМ исходный двоичный информационный поток от источника разделяется на группы, именуемые символами и содержащие i бит. Число всех символов конечно ($P = 2^i$), а их совокупность называется алфавитом. Параметр P определяет позиционность модуляции. Например, если при ЦМ двоичный символ «0» отображается сигналом S_1 , а двоичный символ «1» — S_2 , то $P = 2$, $i = 1$, а модуляция называется двоичной. При использовании такой ЦМ каждый бит от источника передается отдельно, и понятия «бит» и «символ» равнозначны.

В случае $i \neq 1$, $i \in Z$ ЦМ осуществляется посредством формирования символа и дальнейшим выбором одного из $P = 2^i$ детерминированных сигналов для его передачи по каналу связи. Такой вид модуляции называется P -позиционной модуляцией. Например, при $i = 2$ и $P = 4$ символу, содержащему информацию о двух битах, соответствует один из четырех возможных сигналов его представления.

Использование P -позиционной модуляции при $P \neq 2$ приводит к различию в соотношении скорости информационного потока на входе модулятора (информационная скорость, измеряется в бит/с) и на выходе модулятора (техническая скорость, измеряется в бодах/с). Например, при информационной скорости в 128 кбит/с и параметрах модулятора $i = 2$, $P = 4$ техническая скорость 64 кБод (рис. 7.1).

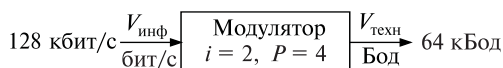


Рис. 7.1. Функционирование модулятора с позиционностью 4

Различная природа используемых сред передачи (физическая проводная линия, радиопередача, волоконно-оптические линии связи) обуславливают различные виды цифровой модуляции (рис. 7.2).

Считается, что исходный цифровой сигнал имеет логический формат двоичных нулей и единиц и с целью передачи должен быть преобразован в сигнал, совместимый с каналом.

В случае *полосовой модуляции* импульсы модулируют синусоиду, называемую *несущей*. Полосовые виды цифровой модуляции позволяют реализовать многоканальную передачу от однотипных источников при использовании одной среды распространения.

В последнее время широкое распространение получили волоконно-оптические линии связи, в которых также используются принципы цифровой модуляции для управления параметрами излучающего лазера.

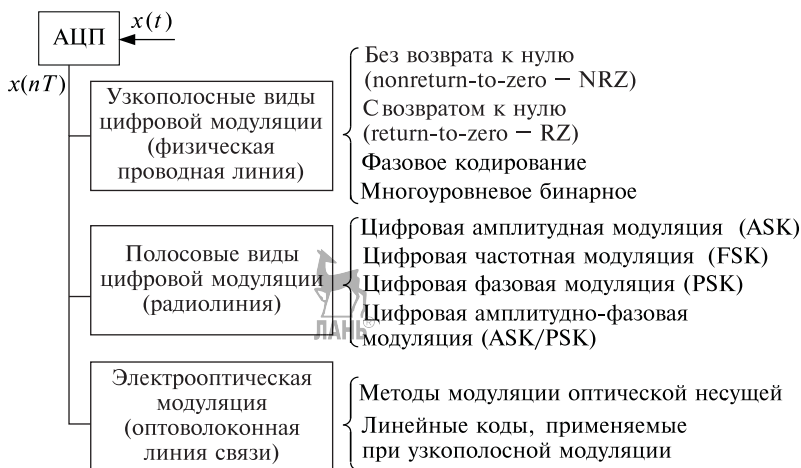


Рис. 7.2. Основные методы цифровой модуляции

Цифровые сообщения, как правило, имеют логический формат двоичных нулей и единиц. Цифры невозможно передать через узкополосный канал, пока они не будут преобразованы в сигналы, совместимые с этим каналом. Для узкополосных каналов такими совместимыми сигналами являются непрерывные импульсные сигналы (видеоимпульсы). ЦМ при этом называют узкополосной (*baseband modulation*). Термин «узкополосный» определяет сигнал, спектр которого начинается от (или около) постоянной составляющей и заканчивается некоторым конечным значением, обычно в несколько мегагерц. В качестве среды распространения при узкополосной ЦМ, как правило, используется физическая проводная линия (кабель).

В случае реализации полосовой модуляции цифровая последовательность модулирует гармоническое колебание, называемое несущей волной или просто несущей.

Полосовые виды цифровой модуляции позволяют реализовать многоканальную передачу от однотипных источников при использовании одной среды распространения.

В последнее время также широкое распространение получили волоконно-оптические линии связи, в которых используются принципы цифровой модуляции для управления параметрами излучающего лазера.

7.1. Полосовая модуляция и демодуляция

Полосовую модуляцию можно определить как процесс варьирования амплитуды, частоты или фазы (или их комбинаций) радио-

частотной несущей согласно передаваемой информации:

$$S(t) = A(t) \cos \Theta(t),$$

где $A(t)$ — переменная во времени амплитуда; $\Theta(t)$ — переменный во времени текущий аргумент несущего колебания. Угол удобно записывать в виде $\Theta(t) = \omega(t)t + \varphi(t)$, где $\omega(t)$ — угловая частота несущей, а $\varphi(t)$ — ее фаза.

При создании систем передачи информации в большинстве случаев оказывается, что спектр исходного сигнала, подлежащего передаче, сосредоточен отнюдь не на тех частотах, которые эффективно пропускает имеющийся канал связи. Кроме того, очень часто необходимо в одном и том же канале связи передавать несколько сигналов одновременно. Одним из способов решения этой задачи является использование частотного разделения каналов, при котором разные сигналы занимают неперекрывающиеся полосы частот.

Во многих случаях требуется, чтобы передаваемый сигнал был узкополосным. Это означает, что эффективная ширина спектра намного меньше его центральной частоты.

Перечисленные причины приводят к необходимости подобной трансформации исходного сигнала, чтобы требования, предъявляемые к занимаемой сигналом полосе частот, были выполнены, а сам исходный сигнал можно было восстановить.

Решение указанной проблемы достигается при использовании модуляции, при которой какой-либо из параметров несущего колебания изменяется во времени пропорционально исходному сигналу. Исходный сигнал называют модулирующим (*modulating signal*), а результирующее колебание с изменяющимися во времени параметрами — модулированным сигналом (*modulated signal*). Обратный процесс — выделение модулирующего сигнала из модулированного колебания — называется демодуляцией (*demodulation*).

При передаче сообщения, представленного дискретными символами, вместо термина «модуляция» применяется термин *манипуляция* (*keying* или *shift keying*), а сам сигнал называется манипулированным.

Основные типы полосовой модуляции/демодуляции представлены в табл. 7.1.

Если при детектировании сигналов приемник использует информацию о фазе несущей, то процесс называется когерентным детектированием, если же подобная информация не используется, процесс именуется некогерентным детектированием. Следует отметить, что термины «детектирование» и «демодуляция» часто в технике цифровой связи используются как синонимы, хотя демодуляция делает

Таблица 7.1

Основные типы цифровой полосовой модуляции

Когерентные схемы	Некогерентные схемы
Фазовая модуляция (PSK) Частотная модуляция (FSK) Амплитудная модуляция (ASK) Модуляция без разрыва фазы (CPM) Смешанные комбинации	Дифференциальная фазовая модуляция (DPSK) Частотная модуляция (FSK) Амплитудная модуляция (ASK) Модуляция без разрыва фазы (CPM) Смешанные комбинации

акцент на восстановлении сигнала, а детектирование — на принятии решения относительно символического значения принятого сигнала.

На рис. 7.3 представлены наиболее характерные методы цифровой полосовой модуляции, к которым относятся фазовая модуляция PSK (phase shift keying), частотная модуляция FSK (frequency shift keying), амплитудная модуляция ASK (amplitude shift keying) и смешанная комбинация ASK и PSK (обозначается как ASK/PSK).

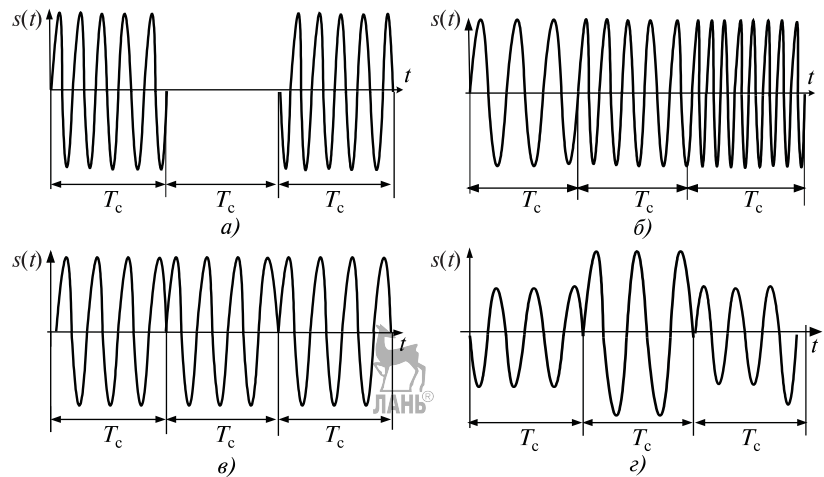


Рис. 7.3. Виды цифровой модуляции: а — ASK; б — FSK; в — PSK; г — ASK/PSK (APK)

Амплитудно-манипулированный сигнал представлен следующими выражениями:

$$S_i(t) = U_i(t) \cos(\omega_0 t + \varphi), \quad i = 1, 2, \dots, P; \quad 0 \leq t \leq T_c,$$

где амплитуда может принимать P значений.

На рис. 7.3 представлен вариант для частного случая ЧМ-сигнала без разрыва фазы. Для частотно-манипулированного сигнала

выражение имеет вид

$$S_i(t) = U \cos(\omega_{0i}t + \varphi), \quad i = 1, 2, \dots, P; \quad 0 \leq t \leq T_c,$$

в котором частота ω_{0i} может принимать P значений.

Цифровая фазовая модуляция (манипуляция) широко используется в коммерческих и военных системах связи. Фазоманипулированный сигнал имеет вид

$$s_i(t) = U \cos(\omega_0 t + 2\pi i/P), \quad i = 1, 2, \dots, P; \quad 0 \leq t \leq T_c,$$

где начальная фаза может принимать P значений.

Наибольший интерес представляют смешанная амплитудно-фазовая модуляция ASK/PSK и ее частный случай, получивший название квадратурной-амплитудной модуляции КАМ (QAM — quadrature amplitude modulation). При использовании данного алгоритма передаваемый сигнал отображается одновременными изменениями амплитуд и начальных фаз синфазных и квадратурных составляющих. Результирующий сигнал S формируется при суммировании этих колебаний. QAM-сигнал может быть представлен соотношением

$$S_m(t) = I_m \cos(2\pi f_c t) + Q_m \sin(2\pi f_c t),$$

где t — время передачи одного символа модуляции; m — номер символа модуляции; f_0 — частота несущей модулируемого колебания; I_m, Q_m — параметры синфазной и квадратурной составляющей модулированного сигнала, учитывающие состояния амплитудного и фазового модулятора для символа.

Этот сигнал может быть представлен в комплексном виде:

$$\dot{S}_m = I_m \cos(\omega_0 t) + jQ_m \sin(\omega_0 t)$$

или

$$\dot{S}_m = A_m \exp j(\omega_0 t + \varphi_m)$$

где $A_m = |\dot{S}_m|$ — амплитуда модулированного сигнала; $\omega_0 = 2\pi f_0$; $(\omega_0 t + \varphi_m) = \text{Arg}(\dot{S}_m)$ — полная мгновенная фаза модулированного сигнала; φ_m — начальная фаза модулированного сигнала.

На рис. 7.4 представлен принцип формирования результирующего колебания S_m путем суммирования вектора квадратурной составляющей Q_m с вектором синфазной составляющей I_m . Амплитуда вектора S определяется соотношением A_m , а угол, который этот вектор образует с осью абсцисс, — соотношением φ_m .

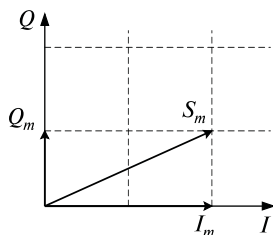


Рис. 7.4. Принцип формирования сигнала QAM

Окончания векторов модулированного колебания образуют прямоугольную сетку на фазовой плоскости действительной — $\text{Re}\{S\}$ и мнимой составляющей вектора модулированного сигнала — $\text{Im}\{S\}$. Число узлов этой сетки определяется типом используемого алгоритма QAM. Схему расположения узлов на фазовой плоскости модулированного QAM колебания принято называть созвездием (constellation). Для указания типа алгоритма QAM принята следующая схема обозначения: QAM- $\langle \text{число} \rangle$, которое обычно представляет собой значение вида 2^i и соответствует количеству узлов на фазовой сетке и максимальному количеству различных значений вектора модулированного сигнала.

На рис. 7.5 приведена структурная схема формирователя модулированного сигнала QAM-16.

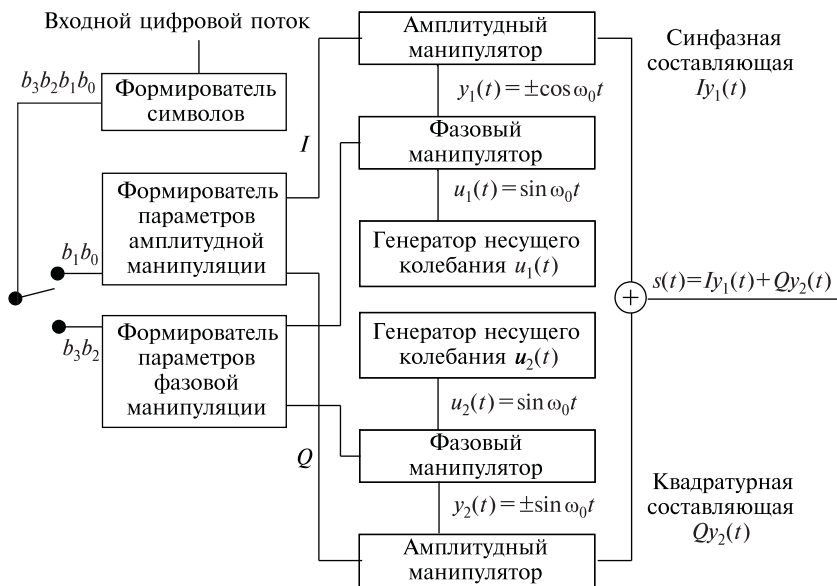


Рис. 7.5. Структурная схема формирователя QAM-сигнала

Выходной сигнал модулятора имеет вид $s(t) = Iy_1(t) + Qy_2(t)$.

Принципы формирования параметров амплитудной и фазовой манипуляций, используемых для формирования синфазной и квадратурной составляющих сигналов QAM-16, представлены в табл. 7.2 и 7.3.

На первом этапе преобразования последовательность бит, которая поступает от источника сигнала, преобразуется с использованием формирователя символов в последовательность $\{b_3, b_2, \dots, b_0\}$.

Таблица 7.2

Принцип формирования параметров
амплитудной манипуляции

b_1	b_0	I	Q
0	0	1	1
0	1	1	3
1	0	3	1
1	1	3	3

Таблица 7.3

Принцип формирования параметров
фазовой манипуляции

b_3	b_2	$y_1(t)$	$y_2(t)$
0	0	$\cos \omega_0 t$	$-\sin \omega_0 t$
0	1	$-\cos \omega_0 t$	$-\sin \omega_0 t$
1	0	$\cos \omega_0 t$	$\sin \omega_0 t$
1	1	$-\cos \omega_0 t$	$\sin \omega_0 t$

Число битов в этом символе равно значению $\log_2 P$ (для алгоритма QAM-16 $i = \log_2 16 = 4$). Формирователь кодовых символов преобразует последовательность $\{b_3, b_2, \dots, b_0\}$ в две последовательности $\{b_1, b_0\}$ и $\{b_3, b_2\}$, которые управляют соответственно амплитудным и фазовым манипуляторами несущих колебаний $\cos \omega_0 t$ и $\sin \omega_0 t$. В соответствии с табл. 7.1 биты $\{b_1, b_0\}$ определяют разрешенные значения амплитуд несущих колебаний $\{1, 3\}$, а в соответствии с табл. 7.2 биты $\{b_3, b_2\}$ определяют начальную фазу несущих колебаний $\{0, \pi\}$. На последнем этапе преобразования выполняются суммирование этих колебаний и формирование результирующего сигнала $S(t)$.

На рис. 7.6 представлено расположение векторов модулированного колебания (созвездие для алгоритма QAM-16). На нем отмечены значения модуляционных символов, которым соответствуют указанные точки на фазовой плоскости модулированного колебания $\{b_3, b_2, b_1, b_0\}$.

Из анализа рис. 7.6 следует, что для алгоритма QAM-16 допустимые значения синфазной и квадратурной составляющих с учетом фазы представлены множеством $\{-3, -1, 1, 3\}$ и определяют значения реальной и мнимой координат вектора модулированного колебания. Так, биты $\{b_3, b_2\}$ определяют номер квадранта фазовой плоскости, а $\{b_1, b_0\}$ — амплитуды несущих колебаний.

Для примера на рис. 7.6 вектором показан сигнал $s_1(t)$, соответствующий передаваемому символу $b_3 b_2 b_1 b_0 = 0010$, при этом используемая нумерация сигнала является условной. В соответствии с табл. 7.2 и 7.3 очевидно, что

$$s_1(t) = 3 \cos \omega_0 t - 1 \sin \omega_0 t. \quad (7.1)$$

На рис. 7.7 представлен график данного сигнала во временной области на длительности двух периодов, полученный в среде MATLAB при частоте $\omega_0 = 2\pi \cdot 2000$ рад/с.

Представление сигнала $s_1(t)$ в гармонической форме с ненулевой

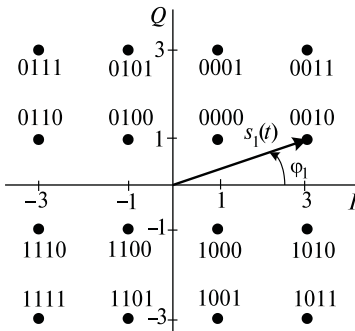


Рис. 7.6. Созвездие сигналов КАМ-16

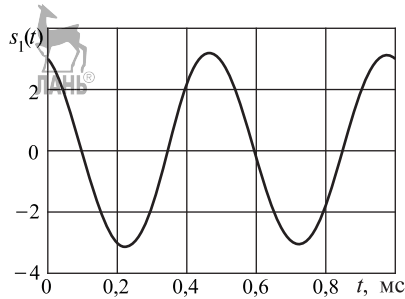


Рис. 7.7. Временное представление сигнала КАМ-16, соответствующего последовательности битов 0010

начальной фазой требует преобразования выражения (7.1):

$$\begin{aligned}
 s_1(t) &= 3 \cos \omega_0 t - 1 \sin \omega_0 t = \sqrt{9+1} \sin \left(\omega_0 t + \operatorname{arctg} \frac{-1}{3} \right) = \\
 &= \sqrt{9+1} \sin \left(\omega_0 t + \arccos \frac{-1}{\sqrt{9+1}} \right) = \\
 &= \sqrt{10} \cos \left(\omega_0 t + \arccos \frac{-1}{\sqrt{9+1}} - \frac{\pi}{2} \right). \quad (7.2)
 \end{aligned}$$

Из выражения (7.2) очевидно, что начальная фаза косинусоидального представления сигнала $s_1(t)$ определяется значением

$$\begin{aligned}
 \varphi_1 &= \arccos \frac{-1}{\sqrt{9+1}} - \frac{\pi}{2} \approx 1,89255 - 1,57080 \approx \\
 &\approx 0,32175 \text{ рад/с} \approx 18,435 \text{ град.}
 \end{aligned}$$

Преобразование модуляционных символов в кодовые символы выполняется с применением алгоритмов Грея, используемого для взаимосвязи помехоустойчивого кодирования данных и отображения модуляционных символов в пространстве Гильберта. Векторам модулированного колебания, которые находятся близко один от другого на фазовой плоскости, ставятся в соответствие значения кодовых символов, которые отличаются значениями только одного бита. В случае ошибки при демодуляции она незначительно повлияет на искажение полезной информации, так как в евклидовом пространстве расстояние между сигналами, представленными кодовыми символами, будет минимально.

Демодулируется сигнал КАМ следующим образом:

- сигнал, полученный из канала связи, умножается на два несущих колебания, которые сдвинуты по фазе относительно друг друга на 90° ;

- результаты умножения пропускаются через ФНЧ с частотой среза, соответствующей значению частоты следования символов модулирующей последовательности;
- на выходе этих ФНЧ будут получены аналоговые сигналы синфазной и квадратурной составляющих;
- эти сигналы дискретизируются с частотой, равной символьной скорости;
- пары отсчетов синфазной и квадратурной составляющих образуют комплексное число, и ближайшая к этому числу точка используемого созвездия (а точнее — соответствующий этой точке информационный символ) выдается в качестве выходного результата.

В настоящее время существует несколько вариантов QAM: алгоритм модуляции QAM-4, который кодирует информационный сигнал изменением фазы несущего колебания с шагом $\pi/2$ (QPSK — *Quadrature Phase Shift Keying*, квадратурная фазовая манипуляция), алгоритмы QAM-16, 32, 64, 128 и 256.

Алгоритм квадратурной амплитудной модуляции является разновидностью алгоритмов амплитудной модуляции и поэтому обладает следующими важными свойствами:

- ширина спектра QAM модулированного колебания не превышает ширину спектра модулирующего сигнала;
- положение спектра QAM модулированного колебания в частотной области определяется номиналом частоты несущего колебания при данном виде модуляции.

Алгоритм квадратурной амплитудной модуляции является относительно простым для реализации и в то же время достаточно эффективным алгоритмом для передачи информации по кабельным линиям связи на основе технологии xDSL (Digital Subscriber Loop).

К недостаткам алгоритма можно отнести относительно невысокий уровень полезного сигнала в спектре модулированного колебания. Этот недостаток является общим для алгоритмов амплитудной модуляции и выражается в том, что максимальную амплитуду в спектре модулированного колебания имеет гармоника с частотой несущего колебания. Данный алгоритм в чистом виде на практике используется достаточно редко.

Гораздо большее распространение получили алгоритмы, которые используют основные принципы QAM и в то же время свободны от его недостатков — это алгоритмы модуляции на основе многомерного представления сигналов и смешанных видов модуляции. Так, цифровая модуляция несущей по амплитуде и фазе позволяет конструировать сигналы, которые соответствуют двумерным векторам

в пространстве диаграмм сигнала (CAP — Carrierless amplitude and phase modulation — амплитудно-фазовая модуляция с подавленной несущей), а если необходимо сконструировать сигнал соответствующий вектору большей размерности, то можно использовать как временную, так и частотную области сигнала для того, чтобы увеличить его размерность и эффективность системы модуляции в целом (OFDM — Orthogonal frequency division multiplexing — ортогональная модуляция с частотным разделением).

Технология OFDM находит широкое применение в протоколах беспроводной связи, например в стандартах IEEE 802.11 и 802.16. Основное преимущество данной технологии заключается в том, что она позволяет реализовать высокую скорость передачи данных, обладает высокой спектральной эффективностью и создает предпосылки для эффективного подавления такого паразитного явления, как многолучевая интерференция сигналов, возникающая в результате многократных отражений сигнала от естественных преград, в результате чего один и тот же сигнал попадает в приемник различными путями. Следовательно, в точке приема результирующий сигнал представляет собой суперпозицию (интерференцию) многих сигналов, имеющих различные амплитуды и смещенных относительно друг друга по времени, что приводит к искажению принимаемого сигнала. Многолучевая интерференция присуща любому типу сигналов, но особенно негативно она сказывается при применении именно широкополосных сигналов.

Для того чтобы избежать многолучевого распространения, в технологии OFDM поток передаваемых данных распределяется по множеству частотных ортогональных друг другу подканалов и передача ведется параллельно на всех подканалах. При этом под ортогональностью каналов подразумевается тот факт, что несущие частоты каждого канала ортогональны друг другу. Ортогональность несущих достигается тем, что несущие частоты каждого подканала выбираются кратными наименьшей частоте для «младшего» подканала, так что в периоде его несущей укладывается целое число периодов несущей другого, т. е. на длительности посылки OFDM-символа укладывается целое число периодов несущих всех подканалов. И хотя сами частотные подканалы могут частично перекрывать друг друга, ортогональность несущих сигналов гарантирует частотную независимость каналов друг от друга, а следовательно, и отсутствие межканальной интерференции.

Для реализации метода OFDM в передающих устройствах используется предварительно мультиплексированный на N каналов сигнал, который модулирует несущие подканалов с использованием

процедур цифровой модуляции на основе алгоритмов ФМ-Р, КАМ-Р. При этом осуществляется временное чередование OFDM-символов по частотным каналам, т. е., например, первый символ передается с использованием OFDM-технологии в первом частотном канале, второй символ — в третьем, третий — во втором, четвертый — опять в первом и т. д. Такой подход позволяет равномерно заполнить весь выделенный для передачи частотный диапазон и снизить влияние многолучевой интерференции на искажение сигнала. Групповой сигнал OFDM формируется путем суперпозиции всех модулированных несущих отдельных подканалов.

Кроме того, для борьбы с многолучевой интерференцией в технологии OFDM используются так называемый охранный интервал (Guard Interval, GI) и циклический префикс (Cycling Prefix). Охранный интервал — это временной промежуток между следованиями отдельных OFDM-символов, необходимый для переключения с одного частотного канала на другой, а циклический префикс добавляется в начало каждого OFDM-символа и представляет собой циклическое повторение окончания символа. Наличие циклического префикса создает временные паузы между отдельными символами, и если длительность охранный интервала превышает максимальное время задержки сигнала в результате многолучевого распространения, то межсимвольной интерференции не возникает.

Циклический префикс является избыточной информацией и в этом смысле снижает полезную (информационную) скорость передачи, но именно он служит защитой от возникновения межсимвольной интерференции. Указанная избыточная информация добавляется к передаваемому символу в передатчике и отбрасывается при приеме символа в приемнике.

В качестве примера высокоскоростной беспроводной технологии соединения устройств, которая базируется на технологии UWB MultiBand OFDM, можно назвать беспроводной интерфейс USB (Wireless USB). Этот новый этап в развитии технологии USB находится в центре внимания группы Wireless USB Promoter Group, в состав которой входят такие отраслевые лидеры, как Agere Systems, HP, Intel, Microsoft, NEC, Philips Semiconductors и Samsung Electronics, а в задачи Wireless USB Promoter Group включена разработка спецификации Wireless USB (WUSB). К настоящему времени уже достигнуты определенные результаты — разработана спецификация WUSB с установленной пропускной способностью 480 Мбит/с. Данная спецификация поддерживает модель использования и архитектуру, которые применяются в проводной технологии USB для высокоскоростного подключения периферийных устройств к хост-компью-

ютуру, что позволяет быстрее перевести сегодняшние проводные USB-решения на новые беспроводные технологии.

7.2. Методы импульсной модуляции

Для передачи исходного цифрового сигнала, представленного в виде последовательности нулей и единиц, по каналу связи часто используется процедура импульсно-кодовой модуляции (ИКМ, pulse-code modulation, PCM). В цифровых системах передачи сигналы ИКМ часто именуются линейными кодами (line code).

Преобразование последовательности нулей и единиц двоичного потока в сигнал ИКМ представлено на рис. 7.8, на котором единица характеризуется наличием импульса, а ноль — его отсутствием (рис. 7.8, а и б). В таком случае сигнал ИКМ (рис. 7.8, в) представляется импульсным сигналом с переменным уровнем, зависящим от чередования нулей и единиц в исходном двоичном потоке.



Рис. 7.8. Вариант представления двоичных цифр в форме сигнала ИКМ

Задачей приемника является определение наличия импульса различного уровня в канале передачи на интервале времени существования информационного бита. Вероятность точного определения наличия импульса является функцией от энергии принятого импульса (или площади под графиком импульса). Следовательно, длительность импульса τ выгодно иметь как можно большую. Если ее увеличить до максимально возможного значения (равного времени пе-

редачи бита), то получится импульсный сигнал, как последовательность переходов между двумя ненулевыми уровнями (биполярное представление). Если сигнал находится на верхнем уровне напряжения, он представляет двоичную единицу, а если на нижнем — двоичный ноль.

При выборе кодировки ИКМ внимание следует обращать на следующие параметры:

1) наличие или отсутствие постоянной составляющей. Удаление из спектра мощностей которой позволяет системе работать на переменном токе. Системы магнитной записи или системы, использующие трансформаторную связь, слабо чувствительны к гармоникам очень низких частот, следовательно, существует вероятность потери низкочастотной информации;

2) помехоустойчивость. Различные типы сигналов ИКМ могут различаться по вероятности появления ошибочных битов при конкретном отношении сигнал/шум. Например, сигналы в кодировке NRZ имеют лучшую достоверность передачи, чем сигналы в униполярной кодировке RZ;

3) возможность автосинхронизации. Структура сигнала (отсутствие длинных серий нулей) должна обеспечивать возможность надежного выделения колебаний тактовой частоты в регенераторах, так, некоторые кодировки ИКМ имеют встроенные функции синхронизации, помогающие восстанавливать синхронизирующий сигнал;

4) структура сигнала должна обеспечивать возможность постоянного контроля достоверности передачи без перерыва связи. Так, некоторые схемы, такие как двубинарная кодировка, предлагают средство выявления информационных ошибок без введения в последовательность данных дополнительных битов выявления ошибок;

5) энергетический спектр цифровых сигналов должен быть сосредоточен в относительно узкой полосе частот при ограничении как по НЧ, так и по ВЧ. Многоуровневые кодировки, повышают эффективность использования полосы, разрешая уменьшение полосы, требуемой для получения заданной скорости передачи данных; следовательно, на единицу полосы приходится больший объем передаваемой информации;

6) возможность дифференциального кодирования. Этот метод позволяет инвертировать полярность сигналов в дифференциальной кодировке, не затрагивая при этом процесс детектирования данных. Это большой плюс в системах связи, в которых иногда происходит инвертирование сигналов;

7) простота реализации приемопередающего оборудования.

Некоторые виды сигналов ИКМ представлены на рис. 7.9. Причина их многообразия заключается в отличии параметров, которые характеризуют каждый код.

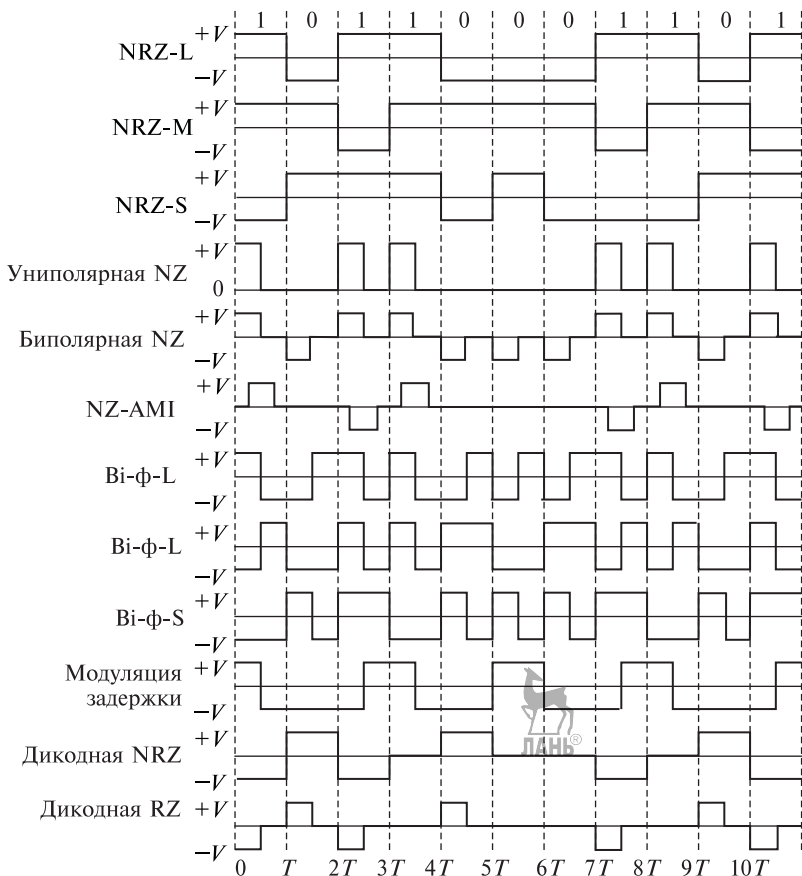


Рис. 7.9. Некоторые виды сигналов ИКМ

Классификация сигналов ИКМ, применяемых к двоичным последовательностям, следующая:

- 1) без возврата к нулю (NRZ — *nonreturn-to-zero*);
- 2) с возвратом к нулю (RZ — *return-to-zero*);
- 3) фазовое кодирование (*bi-φ-X*);
- 4) многоуровневое бинарное кодирование.

Очевидно, что рассмотренная классификация сигналов ИКМ определяет классификационное многообразие соответствующих методов импульсной модуляции.

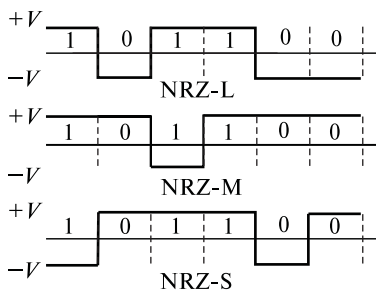


Рис. 7.10. Коды NRZ

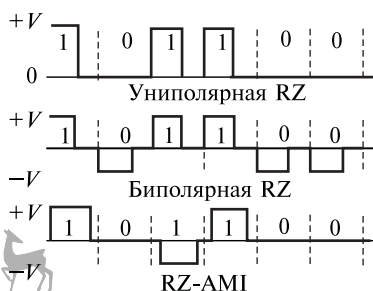


Рис. 7.11. Коды RZ

Группа кодировок NRZ включает следующие подгруппы: NRZ-L (L = level — уровень), NRZ-M (M = mark — метка) и NRZ-S (S = space — пауза) (рис. 7.10).

Код NRZ-L широко используется в цифровых логических схемах. Двоичная единица в этом случае приставляется одним уровнем напряжения, а двоичный нуль — другим. Изменение уровня происходит всякий раз при переходе в последовательности передаваемых битов от нуля к единице или от единицы к нулю.

При использовании кода NRZ-M двоичная единица или метка (mark) представляется изменением уровня, а нуль или пауза (space) — отсутствием изменения уровня. Такая кодировка называется дифференциальной. Применяется код NRZ-M преимущественно при записи на магнитную ленту. Код NRZ-S является обратной к кодировке NRZ-M: двоичная единица представляется отсутствием изменения уровня, а двоичный нуль — изменением уровня.

Группа кодировок RZ включает униполярный код RZ, биполярный код RZ и код RZ-AMI (рис. 7.11). Эти коды применяются при передаче данных и магнитной записи.

В униполярном коде RZ единица представляется наличием импульса, длительность которого составляет половину ширины бита, а нуль — его отсутствием. В биполярном коде RZ единицы и нули представляются импульсами противоположных уровней, длительность каждого из которых также составляет половину ширины бита. В каждом интервале передачи бита присутствует импульс. Код RZ-AMI (AMI — alternate mark inversion — с чередованием полярности) — это схема передачи сигналов, используемая в телефонных системах. Единицы представляются наличием импульсов равных амплитуд с чередующимися полярностями, а нули — отсутствием импульсов.

Группа фазового кодирования включает: *bi-φ-L* (*bi-phase-level* — двухфазный уровень), более известная как манчестерское кодиро-

вание (*Manchester encoding*); *bi-φ-M* (*bi-phase-mark* — метка); *bi-φ-S* (*bi-phase-space* — пауза); модуляция задержки (*delay modulation, DM*), или кодировка Миллера (рис. 7.12). Схемы фазовых кодировок используются в системах магнитной записи и оптической связи, а также в некоторых спутниковых телеметрических каналах передачи данных.

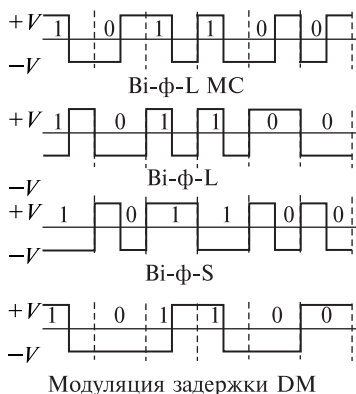


Рис. 7.12. Сигналы фазового кодирования

В коде *bi-φ-L* единица представляется импульсом, длительностью в половину ширины бита, расположенным в первой половине интервала передачи бита, а ноль — таким же импульсом, но расположенным во второй половине интервала передачи бита. В коде *bi-φ-M* в начале каждого интервала передачи бита происходит переход. Единица представляется вторым переходом в середине интервала, ноль — единственным переходом в начале интервала передачи бита. В коде *bi-φ-S* в начале каждого интервала происходит переход. Единица представляется этим единственным переходом, а для представления нуля необходим второй переход в середине интервала. При модуляции задержки единица представляется переходом в середине интервала передачи бита, а ноль — отсутствием иных переходов, если за ним не следует другой ноль.

К кодам, использующим многоуровневое бинарное кодирование, относят те, которые используют три и более уровней. К этой группе относятся сигналы в кодировках биполярной RZ и RZ-AMI. Кроме того, сюда входят схемы, называемые дикодной (*dicode*) и двубинарной кодировкой (*duobinary*).



Рис. 7.13. Дикодные коды

При дикодном коде NRZ переходы в передаваемой информации от единицы к нулю и от нуля к единице меняют полярность импульсов, при отсутствии переходов передается сигнал нулевого уровня (рис. 7.13). При дикодном коде RZ переходы от единицы к нулю и от нуля к единице вызывают изменение полярности длительностью в половину интервала импульса; при отсутствии переходов передается сигнал нуля (см. рис. 7.13).

В настоящее время используются разновидности вышеперечисленных кодов, такие как СМІ (Coded Mark Inversion) — код с инверсией токовых посылок и АМІ (Alternating Mark Inversion) — трехуровневый код с чередованием полярности импульсов.

Код СМІ (рис. 7.14) рекомендован МСЭ-Т для использования в цифровом стыке сигналов четвертого уровня плезиохронной цифровой иерархии (ПЦИ). Энергия низкочастотных составляющих биимпульсных сигналов незначительна и имеет довольно резкий спад составляющих высокой частоты в области $2f_T = 2/\tau_H$. Кроме того, такой сигнал обладает максимальным значением параметра устойчивости признаков тактовой частоты. В коде СМІ единица преобразуется в импульсы с полярностью, противоположной полярности импульса, который соответствовал предыдущей единице, а ноль представляется биимпульсным сигналом определенной фазы. Однако такой сигнал занимает в два раза большую полосу частот по сравнению с исходным униполярным сигналом (по оценке полосы до первого нуля энергетического спектра, в который заключено не менее 90 % энергии), поэтому биимпульсные сигналы используются для передачи цифровых сигналов на относительно небольшие расстояния.

В коде АМІ (рис. 7.14,б) символы нуля исходного двоичного сигнала остаются без изменений, а единицы чередуются по знаку: пос-

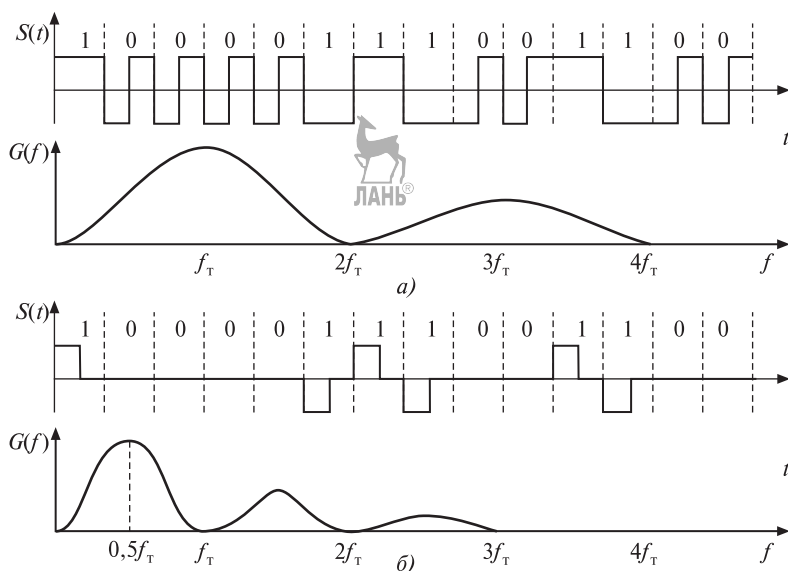


Рис. 7.14. Временное и спектральное представление кодов СМІ и АМІ

ле импульса положительной полярности передается импульс отрицательной полярности и наоборот, положительные и отрицательные импульсы чередуются независимо от числа нулей между ними. Анализ энергетического спектра такого кода показывает, что его сигнал не содержит постоянной составляющей, максимум энергии сосредоточен вблизи частоты $0,5f_T$. Таким образом, важной особенностью сигналов АМІ является то, что для передачи квазитроичного линейного сигнала требуется существенно меньшая полоса частот по сравнению с двоичным (примерно в два раза) и с биимпульсным (в четыре раза) сигналами.

Данные коды не защищены от длинных серий нулей, поэтому он пригоден лишь для малоканальных систем передачи, где вероятность появления многих нулей подряд невелика.

С целью повышения надежности выделения тактовой частоты для большинства сигналов требуется дополнительное преобразование путем изменения статистических свойств, которое осуществляется либо скремблированием, либо путем формирования специальных кодов, повышающих плотность токовых посылок — неалфавитных и алфавитных.

Неалфавитные коды формируются при изменении статистических свойств исходной двоичной информации при некоторых определенных условиях (например, при заданном количестве следующих подряд нулей). Среди них наибольшее распространение получили коды *HDBn* — *High Density Bipolar*, биполярное кодирование с высокой плотностью — в которых последовательность двоичных нулей заданной длины заменяется определенным сочетанием импульсов и пауз (рис. 7.15).

Алгоритм аналогичен АМІ, за исключением кодирования последовательностей более четырех нулей, где каждые четыре нуля за-

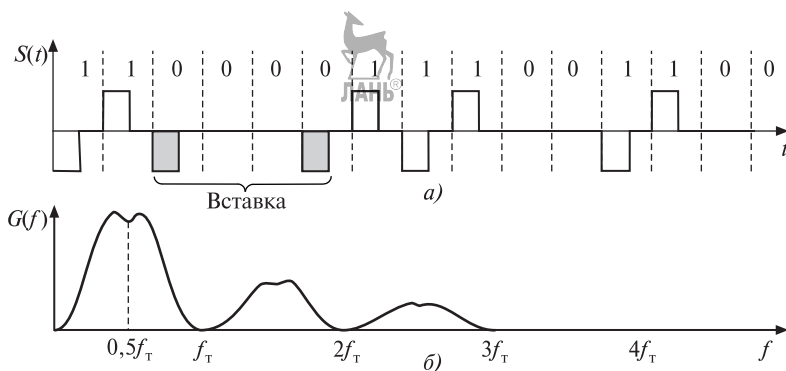


Рис. 7.15. Код HDB-3

меняются последовательностью $000V$, либо $B00V$, где B — импульс по полярности, противоположный предыдущему импульсу (отвечает правилу кодирования АМІ), V (*Violation*) — импульс по полярности, соответствующий предыдущему импульсу (нарушающий правило кодирования АМІ). Замена выбирается таким образом, чтобы, во-первых, число импульсов B между двумя последовательно расположенными импульсами V было нечетным и, во-вторых, чтобы полярности импульсов V чередовались (табл. 7.4).

Таблица 7.4

Правило вставки импульсов замены кода HDB-3

Полярность предыдущего импульса	Число импульсов после последней замены	Правила вставки импульсов
+	Нечетное	000+
—	—	000—
+	Четное	—00—
—	—	+00+

Последовательное изменение знака импульса, нарушающего чередование полярностей в биполярном сигнале, производится с целью выравнивания числа положительных и отрицательных импульсов, т.е. для исключения постоянной составляющей в сигнале. Коды АМІ и HDB-3 позволяет обнаружить ошибки. В коде АМІ ошибка обнаруживается при выявлении нарушения чередования полярности импульсов, а в HDB-3 — вставок. Помеха, приводящая к трансформации символа кодовой последовательности, вызывает нарушение чередования полярности импульсов АМІ или вставок HDB-3.

Код HDB-3 рекомендован стандартом G.703 для цифровых потоков E1, E2 и E3 плезихронной системы иерархии.

Алфавитные коды требуют передачи признаков, достаточных для правильного восстановления границ групп символов кода при декодировании (например, введение специального синхросигнала), что позволяет не только повысить стабильность признаков тактовой частоты, но и увеличить пропускную способность за счет соответствующего снижения значения тактовой частоты. Для алфавитных кодов приняты следующие обозначения: $nBkM$ (n указывает число символов в кодируемой группе исходной двоичной последовательности, B (*Binary*) — двоичное основание счисления, используемое в исходной последовательности, k — число символов в группе кода, в который преобразовывается исходная комбинация, а последняя буква (или сочетание букв) отражает новое основание счисления: T — троичное (*ternary*); Q — четверичное (*quaternary*) и т.д.) Например,

обозначение $3B2T$ указывает, что каждая группа из трех двоичных символов преобразуется в соответствующую группу из двух троичных символов.

Технология *HDSL (High-bit-rate Digital Subscriber Loop)* основывается на использовании двух видов цифровой модуляции — $2B1Q$ (*2 Binary 1 Quarternary*) и *CAP (Carrierless Amplitude and Phase Modulation)*.

Для формирования линейного кода входной информационный поток делится на кодовые группы по два бита в каждой. В зависимости от комбинации значений битов кодовой группы ей ставится в соответствие один из четырех кодовых символов, каждому из которых, в свою очередь, ставится в соответствие один из уровней кодового напряжения (табл. 7.5 и рис. 7.16).

Таблица 7.5
Правила формирования кода $2B1Q$

Кодовая группа	Кодовый символ	Кодовое напряжение, В
00	−3	−2,5
01	−1	−0,833
10	+3	+2,5
11	+1	+0,833

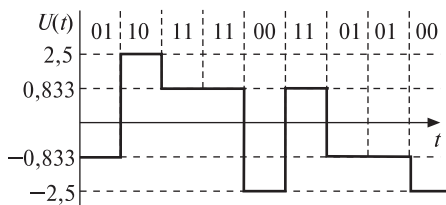


Рис. 7.16. Код $2B1Q$

Достоинством данного алгоритма является простота и дешевизна его реализации. К недостаткам этого метода линейного кодирования следует отнести невысокую спектральную эффективность и, следовательно, ограниченные возможности для передачи информационного сигнала по зашумленным линиям с большим затуханием, а также необеспечение поддержки баланса положительных и отрицательных импульсов выходного напряжения и, следовательно, код $2B1Q$ должен быть предварительно обработан специальными процедурами, которые должны обеспечить подавление постоянной составляющей (скремблирование).

7.3. Модуляция для волоконно-оптических линий связи

При передаче цифрового сигнала по оптическому волокну предъявляют определенные требования к свойствам такого сигнала.

Требование по ограничению энергетического спектра. Ограничение спектра в области низких частот вызвано требованиями безыскаженной передачи в усилителе фотоприемника, имеющего цепи развязки по постоянному току. Другой причиной ограничения спектра снизу является необходимость стабилизации выходной

мощности лазерного диода передающего оптического модуля (мощность лазерного диода может быть стабилизирована введением отрицательной обратной связи по среднему значению излучаемой мощности только в том случае, если подавлена низкочастотная часть спектра, изменяющаяся во времени под действием информационного сигнала). Ограничение спектра в области ВЧ снижает мощность помех, а также уровень межсимвольной помехи.

Требование по ограничению числа уровней в линейном сигнале. При реализации цифровых систем передачи в общем случае устанавливается P возможных значений уровня цифрового сигнала. Так, в соответствии с Рекомендацией МСЭ-Т G.703, для 1–3 ступеней иерархии цифровых систем передачи установлен код HDB-3 с чередующейся полярностью импульсов. В волоконно-оптических линиях связи импульсные посылки излучаемой оптической мощности могут быть только положительными или нулевыми, поскольку интенсивность оптического излучения по самой природе является положительной величиной. Поэтому непосредственное использование биполярных кодов типа HDB-3 невозможно.

Кроме того, применение многоуровневых кодов (даже положительной полярности) сопряжено с определенными техническими трудностями: во-первых, присутствуют нелинейность модуляционной характеристики и температурная зависимость мощности излучателя; во-вторых, в термехуровневом коде мощность сравниваемых символов при принятии решения оказывается в 2 раза меньше, что эквивалентно уменьшению мощности передатчика в 2 раза.

Очень широкое применение нашли в ВОЛС коды группы NRZ, которые используются для передачи синхронных транспортных модулей (STM).

Однако коды класса NRZ или RZ не повышают скорости передачи в линии. Общим недостатком этих кодов является то, что они не удовлетворяют большинству перечисленных выше требований (спектр сигнала сосредоточен в низкочастотной области, имеются значительная постоянная составляющая и т. д.). Такие коды могут применяться на коротких участках без промежуточных регенераторов. Для снижения в спектре низкочастотных составляющих применяют бифазный BIF (манчестерский) код. При этом отсутствуют более двух следующих подряд идентичных символов, что и позволяет снизить содержание в спектре низкочастотных составляющих. Основные недостатки: низкая помехозащищенность, сложности с выделением тактовой частоты и детектированием ошибок.

На рис. 7.17 представлены коды, применяемые в ВОЛС.

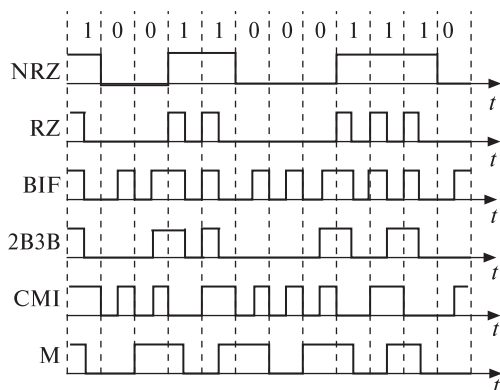


Рис. 7.17. Линейные коды ВОЛС

В относительно низкоскоростных волоконно-оптических системах передачи (до второй ступени иерархии) применяется код СМІ, в котором каждому двоичному символу исходного сигнала сопоставляются два двоичных символа кода в линии (символу «0» исходного сигнала ставится в соответствие последовательность символов «10», а символу «1» — попеременно последовательности «11» и «00»).

В системах более высоких ступеней иерархии применяется код Миллера, в котором нулю исходной последовательности ставятся в соответствие чередующиеся кодовые слова «11» или «00», а единице — «01» или «10». При этом соседние переходы вида 10 и 01 в линейном сигнале будут отстоять друг от друга не ближе, чем на тактовый интервал T , и не дальше, чем на $2T$. Поэтому основная часть энергетического спектра сосредоточена в области ниже тактовой частоты, а низкочастотная составляющая оказывается частично подавленной (30 % от низкочастотной составляющей NRZ). Контроль ошибок ведется при появлении переходов с частотой, большей $1/T$.

Для третичной и четверичной систем используются коды класса $mBnB$ ($m \geq 2$). В кодах этого класса последовательность исходного сигнала разбивается на блоки из m бит, каждый из которых преобразуется в блок из n символов в соответствии с определенным алфавитом. Алфавит выбирается с таким расчетом, чтобы уменьшить максимальное число последовательных одинаковых символов и спектральную плотность сигнала в низкочастотной области.

В отдельную группу выделяются коды со вставками. При формировании таких кодов к информационным символам добавляется один дополнительный символ-вставка C ($mB1C$), формируемый в зависимости от значения последнего информационного символа в

блоке из m импульсов, или D ($mB1D$), формируемый в зависимости от количества символов «1» в блоке из m импульсов, например $3B1C$, $8B1C$, $10B1D$. К этому же классу относятся коды $mB1D1R$, в которых помимо вставки D к блоку из m символов добавляется один бит R для служебной связи.

Таким образом, выбор кода в линии является достаточно сложной и важной проблемой, правильное решение которой во многом определяет технико-экономические показатели ЦСП в целом.



Заключение

Очевидно, что круг вопросов, рассмотренных в данном пособии, не охватывает всего многообразия существующих и перспективных направлений, методов и средств ЦОС. Это связано как с ограниченным объемом курса, так и с методической сложностью изучения и отражения в учебном издании новых научных результатов данной отрасли знаний.

Становление и развитие специалиста в области цифровой связи невозможны без постоянного ознакомления и изучения новых научных результатов и современных средств освоения и реализации систем ЦОС, которыми являются универсальные пакеты программ для научных исследований и автоматизированного проектирования — MatLab, LabVIEW, SystemView, Code Composer studio, MC Studio, Quartus.

В настоящее время ЦОС как наука активно развивается, и последние ее достижения преимущественно излагаются в научной литературе, монографиях, трудах конференций, периодических изданиях, в издаваемом специализированном журнале «Цифровая обработка сигналов» (URL: www.dsps.ru), а также на сайтах Интернета. Поэтому углубленное изучение цифровой обработки сигналов невозможно без самостоятельного обращения к данным источникам информации.

Следует подчеркнуть, что знание основ цифровой обработки сигналов является в настоящее время востребованным при работе в любой из областей профессиональной деятельности, использующих информационные технологии.

Основные сокращения

АДИКМ	— адаптивная дифференциальная импульсно-кодовая модуляция
АКФ	— автокорреляционная функция
АР	— авторегрессия
АРСС	— авторегрессия со скользящим средним
АЦФ	— адаптивный цифровой фильтр
АЦП	— аналого-цифровое преобразование
АЧХ	— амплитудно-частотная характеристика
БИХ	— бесконечная импульсная характеристика
БПФ	— быстрое преобразование Фурье
ВОЛС	— волоконно-оптические линии связи
ВОСП	— волоконно-оптические системы передачи
ГВЗ	— групповое время задержки
ДИКМ	— дифференциальная импульсно-кодовая модуляция
ДКП	— дискретное косинусное преобразование
ДПФ	— дискретное преобразование Фурье
ДПЛ	— дискретное преобразование Лапласа
ДС	— дискретная система
ЗС	— звуковой сигнал
ИКМ	— импульсно-кодовая модуляция
ИХ	— импульсная характеристика
КАМ	— квадратурная амплитудная модуляция
КИХ	— конечная импульсная характеристика
КС	— канал связи
КТЧ	— канал тональной частоты
ЛДС	— линейная дискретная система
ЛП	— линейное предсказание
ЛРУ	— линейное разностное уравнение
МСЭ	— Международный союз электросвязи
МСЭ-Т	— Сектор стандартизации в области телекоммуникаций МСЭ
НРЦФ	— нерекурсивный цифровой фильтр
ОДПФ	— обратное дискретное преобразование Фурье
ОТ	— основной тон
ОФ	— оконная функция
ПДПФ	— прямое дискретное преобразование Фурье
ПЛИС	— программируемая логическая интегральная схема
ПФ	— передаточная функция, полосовой фильтр
ПХ	— переходная характеристика
РС	— речевой сигнал
РУ	— решающее устройство
РЦФ	— рекурсивный цифровой фильтр
РФ	— режекторный фильтр
СнК	— система на кристалле
СПМ	— спектральная плотность мощности
СС	— скользящее среднее

УДО	— устройство дополнительной обработки
УПЦФ	— узкополосный цифровой фильтр
УСУ	— устройство сравнения и управления
ФВЧ	— фильтр верхних частот
ФНЧ	— фильтр нижних частот
ФЧХ	— фазочастотная характеристика
ЦОС	— цифровая обработка сигналов
ЦСА	— цифровой спектральный анализ
ЦФ	— цифровой фильтр
ЧХ	— частотная характеристика
ЦАП	— цифро-аналоговое преобразование
ЦЛТ	— цифровой линейный тракт
ЦМ	— цифровой модулятор
ЦОС	— цифровая обработка сигналов
ЦПОС	— цифровой процессор обработки сигналов
ЦС	— цифровой сигнал
ЦСП	— цифровая система передачи
ЧПИ	— чередование полярности импульсов
ШРС	— широкополосный речевой сигнал
ЭШП	— эквивалентная шумовая полоса
ACELP	— линейное предсказание с возбуждением от алгебраического кода (Algebraic Code Excited Linear Prediction)
ADPCM	— адаптивная дифференциальная импульсно-кодовая модуляция (Adaptive Differential Pulse Code Modulation)
AMR-WB	— адаптивный многоскоростной кодек широкополосного речевого сигнала (Adaptive Multi-Rate Wideband speech codec)
CELP	— линейное предсказание с возбуждением от кода (Code Excited Linear Prediction)
AR	— авторегрессия (Autoregressive)
ARMA	— авторегрессия со скользящим средним (Autoregressive Moving Average)
DPCM	— дифференциальная импульсно-кодовая модуляция (Differential Pulse Code Modulation)
EV	— собственные векторы (Eigenvectors)
ITU	— Международный союз электросвязи (International Telecommunication Union) — специализированное учреждение Организации Объединенных Наций
ITU-T	— сектор стандартизации в области телекоммуникаций МСЭ (Telecommunication Standardization Sector of ITU)
FFT	— быстрое преобразование Фурье (Fast Fourier Transform)
MA	— скользящее среднее (Moving Average)
MBSA	— модельный спектральный анализ (Model-Based Spectrum Analysis)
MUSIC	— классификация множественных сигналов (Multiple Signal Classification)
QMF USB	— квадратурный зеркальный фильтр (Quadrature Mirror Filter) — Universal Serial Bus



Условные обозначения

$A(\omega T)$ — амплитудно-частотная характеристика дискретного фильтра (дискретной системы)

$A(\omega)$ — амплитудно-частотная характеристика аналогового фильтра (аналоговой системы)

A — амплитуда гармонического колебания

a_m — масштабирующий коэффициент множителя m -й прямой ветви дискретного фильтра (дискретной системы)

$a_{\text{ш}}$ — коэффициент включения белого шума

$B(m)$ — автокорреляционная функция

b_k — масштабирующий коэффициент множителя k -й обратной ветви дискретного фильтра (дискретной системы)

D — коэффициент перекрытия соседних интервалов наблюдения при спектральном оценивании

E — среднеквадратическая ошибка линейного предсказания

$e(n)$ — сигнал ошибки линейного предсказания

F — оператор преобразования цифрового устройства

$F[x(nT)]$ — спектр дискретного сигнала $x(nT)$

$\Delta F_{\text{ш}}$ — эквивалентная шумовая полоса оконной функции

f — частота

f_d — частота дискретизации

f_0 — нормированная частота

Δf — частотное разрешение спектрального оценивания

$G(\omega)$ — спектральная плотность мощности

$G_p(\omega)$ — распределение мощности реализации сигнала

$g(n), g(nT)$ — переходная характеристика дискретного фильтра (дискретной системы)

$H(e^{j\omega T})$ — частотная характеристика дискретного фильтра (дискретной системы)

$H(z)$ — передаточная функция дискретного фильтра (дискретной системы)

$h(n), h(nT)$ — импульсная характеристика дискретного фильтра (дискретной системы)

$K - 1$ — порядок нерекурсивного дискретного фильтра (дискретной системы), степень полинома числителя передаточной функции

k — номер прямой ветви дискретного фильтра (дискретной системы), $k = 0, 1, \dots, K - 1$; номер частотной составляющей в ДПФ

L — общее число отсчетов сигнала, используемое при спектральном оценивании сигнала

M_6 — максимальный уровень боковых лепестков оконной функции

$M - 1$ — порядок рекурсивного дискретного фильтра (дискретной системы), степень полинома знаменателя передаточной функции, порядок линейного предсказания

m — номер обратной ветви дискретного фильтра (дискретной системы), $m = 0, 1, \dots, M - 1$; аргумент АКФ

N — количество отсчетов на периоде дискретного сигнала, количество спектральных составляющих на периоде дискретного преобразования Фурье (ДПФ)

- n — порядковый номер отсчета, $n = 0, 1, 2, 3, \dots$
 p — комплексная переменная преобразования Лапласа, $p = \sigma + j\omega$
 $p(n), p(nT)$ — оконная функция (функция взвешивания)
 R — радиус окружности в z -области, являющейся границей области сходимости z -преобразования
 T — период дискретизации
 T_a — период аналогового сигнала $x(t)$
 $T(p)$ — операторная передаточная функция аналогового фильтра (аналоговой системы)
 $T(j\omega)$ — комплексная передаточная функция (частотная характеристика) аналогового фильтра (аналоговой системы)
 t_n — длительность интервала наблюдения
 $\text{test}(n)$ — тестовый дискретный сигнал
 $u(n), u(nT)$ — дискретный белый шум
 $u_1(n), u_1(nT)$ — цифровой единичный скачок
 $u_3(n), u_3(nT)$ — дискретный экспоненциальный сигнал
 V_6 — асимптотическая скорость спада боковых лепестков оконной функции
 W_N — поворачивающий множитель ДПФ
 $X(j\omega)$ — спектр аналогового непериодического сигнала $x(t)$
 $X(e^{j\omega T}), X(j\omega T)$ — спектр дискретного сигнала $x(nT)$
 $X(jk\Omega), X(jk)$ — дискретное преобразование Фурье периодического сигнала $x(nT)$ (сигнала конечной длительности), $k = 0, 1, 2, \dots, N - 1$
 $X(jk\Omega_a)$ — спектр аналогового периодического сигнала $x(t)$, $k = 0, 1, 2, \dots$
 $X(p)$ — p -изображение сигнала (прямое преобразование Лапласа)
 $X(z)$ — z -изображение дискретного сигнала $x(nT)$
 $x(t)$ — аналоговый сигнал
 $x(n), x(nT)$ — дискретный сигнал
 $x_p(nT)$ — взвешенный дискретный сигнал
 $x_n(n), x_n(nT)$ — цифровой сигнал
 $Z[x(nT)]$ — z -изображение дискретного сигнала $x(nT)$
 $Z^{-1}[H(z)]$ — обратное z -преобразование
 z — комплексная переменная дискретного z -преобразования, $z = \alpha + j\beta = re^{j\varphi}$
 z_i^0 — нуль z -изображения дискретного сигнала, $i = 1, 2, 3, \dots$
 z_i^* — полюс z -изображения дискретного сигнала, $i = 1, 2, 3, \dots$
 z_k^0 — нуль передаточной функции дискретной системы (цифрового фильтра), $k = 1, 2, \dots, K - 1$
 z_m^* — полюс передаточной функции дискретной системы (цифрового фильтра), $m = 1, 2, \dots, M - 1$
 $\delta(n), \delta(nT)$ — цифровой единичный импульс
 ν — номер интервала наблюдения
 σ^2 — дисперсия случайного сигнала
 $\tau(\omega T)$ — групповое время замедления дискретного фильтра (дискретной системы)
 φ — начальная фаза гармонического колебания; фазовый угол комплексной переменной z
 $\varphi(\omega)$ — фазочастотная характеристика аналогового фильтра (аналоговой системы)
 $\varphi(\omega T)$ — фазочастотная характеристика дискретного фильтра (дискретной системы)
 $\varphi(n), \varphi(nT)$ — базисная функция преобразования Фурье непериодического дискретного сигнала

$\varphi_k(n), \varphi_k(nT)$ — базисная функция дискретного преобразования Фурье, $k = 0, 1, 2, \dots, N - 1$

Ω — основная угловая частота дискретного преобразования Фурье

Ω_d — угловая частота первой гармоники спектра непрерывного периодического сигнала

ω — угловая частота

ω_d — угловая частота дискретизации

ω_0 — нормированная угловая частота



Литература



1. **Аоки М.** Введение в методы оптимизации. — М.: Наука, 1977. — 334 с.
2. **Айфичер Э.С., Джервис Б.У.** Цифровая обработка сигналов: практический подход. 2-е изд. — М.: Вильямс, 2004. — 992 с.
3. **Улахович Д.А.** Основы теории линейных электрических цепей: учебное пособие. — СПб.: БХВ-Петербург, 2009. — 816 с.
4. **Блейхут Р.** Быстрые алгоритмы цифровой обработки сигналов. — М.: Мир, 1989. — 448 с.
5. **Быков С.Ф., Журавлев В.И., Шалимов И.А.** Цифровая телефония: учебное пособие для вузов. — М.: Радио и связь, 2003. — 144 с.
6. **Гольденберг Л.М., Матюшкин Б.Д., Поляк М.Н.** Цифровая обработка сигналов: учеб. пособие для вузов. — М.: Радио и связь, 1990. — 256 с.
7. **Гоноровский И.С.** Радиотехнические цепи и сигналы: учеб. для вузов. — М.: Радио и связь, 1994. — 608 с.
8. **Гонсалес Р., Вудс Р.** Цифровая обработка изображений. — М.: Техносфера, 2005. — 1072 с.
9. **Каппелини В., Константи́нидис Дж, Эмилиани П.** Цифровые фильтры и их применение. — М.: Энергоатомиздат, 1983. — 360 с.
10. **Корн Г., Корн Т.** Справочник по математике. — М.: Наука, 1984. — 832 с.
11. **Крошьер Р.Е., Рабинер Л.Р.** Интерполяция и децимация цифровых сигналов: метод, обзор // ТИИЭР. 1981. Т. 69, № 3. С. 14–40.
12. **Куприянов М.С., Матюшкин Б.Д.** Цифровая обработка сигналов: процессоры, алгоритмы, средства проектирования. СПб.: Политехника, 1998. — 592 с.
13. **Ланнэ А.А., Матюшкин Б.Д., Улахович Д.А.** Основы ЦОС. — СПб.: ВАС, 1995. — 132 с.
14. **Маркел, Дж.Д., Грей А.Х.** Линейное предсказание речи. — М.: Связь, 1980. — 307 с.
15. **Марпл С.Л. (мл.)** Цифровой спектральный анализ и его приложения. Пер. с англ. — М.: Мир, 1990. — 584 с.
16. **Нуссбаумер Г.** Быстрое преобразование Фурье и алгоритмы вычисления свертки. — М.: Радио и связь, 1985. — 248 с.
17. **Оппенгейм А., Шафер Р.** Цифровая обработка сигналов. — М.: Связь, 1979. — 416 с.
18. **Оппенгейм А.** Применение цифровой обработки сигналов. — М.: Мир, 1980. — 552 с.
19. **Рабинер Л., Гоулд Б.** Теория и применение цифровой обработки сигналов. — М.: Мир, 1978. — 848 с.
20. **Сергиенко А.Б.** Цифровая обработка сигналов: учебное пособие для вузов. — СПб.: Питер, 2002. — 608 с.
21. **Солонина А.И., Улахович Д.А., Яковлев Л.А.** Алгоритмы и процессоры цифровой обработки сигналов: учебное пособие. — СПб.: БХВ-Петербург, 2001. — 464 с.
22. **Солонина А.И., Улахович Д.А., Арбузов С.М. и др.** Основы цифровой обработки сигналов: курс лекций. — СПб.: БХВ-Петербург, 2003. — 608 с.
23. **Уидроу В., Стирнз С.** Адаптивная обработка сигналов. — М.: Радио и связь, 1989. — 440 с.

Оглавление

Введение.....	3
Глава 1. Дискретные и цифровые сигналы	6
1.1. Предмет и проблематика цифровой обработки сигналов	6
1.2. Функциональная схема системы ЦОС. Достоинства и недостатки ЦОС	8
1.3. Классификация сигналов и их математическое описание	10
1.4. Квантование чисел в ЦОС.....	13
1.5. Анализ шумов квантования сигналов.....	14
1.6. Представление и кодирование чисел в ЦОС	18
1.7. Типовые дискретные сигналы	24
1.8. Основные операции цифровой обработки сигналов ...	26
1.9. Спектры непериодических дискретных сигналов и их свойства	29
1.10. Соотношение спектров непериодических аналоговых и дискретных сигналов	32
1.11. Спектры периодических дискретных сигналов	35
1.12. Характеристика спектрального представления различных классов сигналов на основе преобразований Фурье	43
1.13. Быстрое преобразование Фурье.....	45
1.13.1. Быстрое преобразование Фурье с прореживанием по времени.....	46
1.13.2. Быстрое преобразование Фурье с прореживанием по частоте	52
1.14. Дискретное косинусное преобразование	57
1.15. Дискретное преобразование Лапласа. Z-преобразование дискретных сигналов и его свойства	59
1.16. Связь между преобразованиями Лапласа, Фурье и z-преобразованием дискретных сигналов	63
1.17. Вейвлет-преобразование	65
1.17.1. Особенности применения и использования вейвлет-анализа.....	65
1.17.2. Непрерывный вейвлет-анализ.....	68
1.17.3. Дискретный вейвлет-анализ	72

Глава 2. Основы цифровой фильтрации	76
2.1. Дискретные системы	76
2.2. Принцип цифровой фильтрации	78
2.3. Математические модели и характеристики цифровых фильтров	79
2.4. Классификация цифровых фильтров	89
2.5. Структурные схемы фильтров	91
2.6. Простейшие цифровые фильтры и их характеристики	96
2.7. Расчет характеристик рекурсивного цифрового фильтра во временной, частотной и z -областях	109
2.8. Проектирование цифровых фильтров с конечной импульсной характеристикой	114
2.9. Элементы проектирования цифровых фильтров с бесконечной импульсной характеристикой	118
2.10. Специальные цифровые фильтры	122
2.10.1. Цифровые преобразователи Гильберта	123
2.10.2. Цифровые дифференциаторы и интеграторы	128
2.10.3. Цифровые согласованные КИХ-фильтры	131
2.11. Собственные шумы цифровых устройств	133
2.12. Адаптивные цифровые фильтры	139
Глава 3. Цифровая многоскоростная обработка сигналов	145
3.1. Методы преобразования частоты. Классификация систем многоскоростной цифровой обработки сигналов	145
3.2. Однократные системы интерполяции и децимации	146
3.2.1. Однократная система интерполяции	146
3.2.2. Однократная система децимации	151
3.3. Полифазные структуры систем интерполяции и децимации	155
Глава 4. Цифровой спектральный анализ	168
4.1. Задачи и особенности цифрового спектрального анализа	168
4.2. Классификация методов цифрового спектрального анализа	172
4.3. Непараметрический цифровой спектральный анализ	173
4.3.1. Цифровой спектральный анализ методами цифровой фильтрации	174
4.3.2. Цифровое спектральное оценивание на основе дискретного преобразования Фурье	175
4.3.3. Оконные функции	178

4.3.4. Цифровое спектральное оценивание методом усреднения модифицированных периодограмм	185
4.4. Параметрический цифровой спектральный анализ ..	187
4.4.1. Принцип параметрического спектрального оценивания	188
4.4.2. Параметрическое спектральное оценивание на основе моделей сигналов в виде реакции линейных цифровых фильтров	189
4.4.3. Параметрическое спектральное оценивание на основе модели авторегрессионного процесса	191
4.4.4. Параметрическое спектральное оценивание на основе модели процесса скользящего среднего	197
4.4.5. Параметрическое спектральное оценивание на основе модели процесса авторегрессии — скользящего среднего ..	198
Глава 5. Цифровая обработка речевого сигнала	200
5.1. Процесс речеобразования	200
5.2. Характеристики речевого сигнала	203
5.3. Метод линейного предсказания дискретного речевого сигнала	208
5.4. Эффективное кодирование речевого сигнала	213
5.4.1. Задачи, параметры и классификация кодеков речи ..	213
5.4.2. Методы кодирования формы волны речевого сигнала	216
5.4.3. Вокодерное кодирование	224
5.5. Гибридное кодирование речевого сигнала на основе метода линейного предсказания	227
5.5.1. Основные процедуры гибридного кодирования речевого сигнала	227
5.5.2. Кодирование речевого сигнала в системе связи стандарта GSM	237
5.5.3. Метод анализа через синтез	240
5.5.4. Кодирование речевых сигналов в системе связи стандарта TETRA	241
5.5.5. Стандарты G.728 и G.729	244
5.6. Кодирование широкополосного речевого сигнала	252
5.6.1. Стандарт G.722	253
5.6.2. Стандарт G.722.2	257
5.7. Оценка качества передачи речевого сигнала	263
Глава 6. Цифровая обработка изображений	269
6.1. Математические модели изображений	269
6.2. Описание получателя изображений	271
6.3. Стандарты кодирования изображений	273
6.3.1. Стандарты кодирования неподвижных изображений ..	273
6.3.2. Форматы представления неподвижного изображения ..	293

6.3.3. Принципы и особенности кодирования сигнала подвижного изображения	301
6.3.4. Стандарты кодирования сигнала подвижного изображения серий H.26x и MPEG	308
Глава 7. Цифровая модуляция и демодуляция	322
7.1. Полосовая модуляция и демодуляция	324
7.2. Узкополосная (импульсная) модуляция и демодуляция	334
7.3. Модуляция для волоконно-оптических линий связи .	342
Заключение	346
Основные сокращения	347
Условные обозначения	349
Литература	352

