

Министерство образования Российской Федерации

**ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ СИСТЕМ
УПРАВЛЕНИЯ И РАДИОЭЛЕКТРОНИКИ (ТУСУР)**

Кафедра телевидения и управления

М.И. Курячий

ЦИФРОВАЯ ОБРАБОТКА СИГНАЛОВ

Учебное пособие

2002

Рецензент: к. т. н., доцент кафедры ТУ ТУСУР Кормилин В.А.

Корректор: Красовская Е.Н.

Курячий М.И.

Цифровая обработка сигналов: Учебное пособие. – Томск: Томский межвузовский центр дистанционного образования, 2002. – 175 с.

В учебном пособии изложены основы цифровой обработки одномерных и многомерных сигналов. Подробно изложены вопросы анализа и синтеза цифровых фильтров, включая расчёт точностных характеристик и оценку погрешностей округления цифровых данных. В учебное пособие также включён материал по методам цифровой обработки изображений и их практическому использованию. Данное пособие в сочетании с разработанным компьютерным лабораторным практикумом и набором заданий позволяет с большой степенью самостоятельной работы студентам освоить данную дисциплину.

Пособие предназначено для студентов специальностей радиотехнического профиля.

© Курячий Михаил Иванович, 2002

© Томский межвузовский центр
дистанционного образования, 2002

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	6
1. АНАЛИЗ ЦИФРОВЫХ ФИЛЬТРОВ	11
1.1. Цифровые цепи и сигналы	11
1.1.1. Цифровые сигналы.....	11
1.1.2. Аналого-цифровые и цифро-аналоговые преобразователи.....	13
1.1.3. Цифровые системы обработки сигналов	18
1.2. Дискретное преобразование Фурье (ДПФ).....	22
1.3. Z-преобразование.....	23
1.4. Алгоритмы функционирования и формы реализации линейных цифровых фильтров	26
1.5. Характеристики ЛЦФ с постоянными параметрами	30
1.5.1. Системная (передаточная) функция фильтра в Z-форме.....	30
1.5.2. Импульсная и переходная характеристики.....	31
1.5.3. Частотные характеристики цифровых фильтров	33
1.5.4. Групповое время запаздывания.....	35
1.5.5. Устойчивость цифровых фильтров.....	35
1.5.6. Точностные характеристики цифровых фильтров.....	40
1.7. Нелинейные эффекты в цифровых фильтрах	52
1.7.1. Эффекты квантования	52
1.7.2. Ошибки квантования в ЦРФ.....	54
1.7.3. Методы борьбы с нелинейными эффектами в рекурсивных ЦФ	56
1.7.4. Точность и эффективность цифровых вычислений с сохранением остатков	60
1.7.5. Особенности построения каналов слежения с использованием ЦФ	65
2. СИНТЕЗ ЦИФРОВЫХ ФИЛЬТРОВ	69
2.1. Синтез цифровых устройств для обработки одномерных данных	69
2.2. Синтез ЦФ по методу инвариантного преобразования импульсной характеристики	70
2.3. Синтез цифровых фильтров методом отображения дифференциалов.....	78
2.4. Синтез цифровых фильтров методом билинейного преобразования.....	81
2.5. Метод синтеза ЦФ с использованием Z-форм.....	88
2.6. Частотные преобразования, применяемые при синтезе цифровых фильтров.....	90
2.7. Прямой синтез ЦФ.....	95
2.8. Методы синтеза фильтров с КИХ.....	96
2.8.1. Метод частотной выборки	96
2.8.2. Метод временных окон	99

3. ЦИФРОВАЯ ОБРАБОТКА ИЗОБРАЖЕНИЙ	106
3.1. Представление и преобразование двумерных сигналов.....	109
3.1.1. Некоторые особые двумерные последовательности.....	110
3.1.2. Многомерные системы.....	111
3.1.3. Базовые операции, используемые в многомерных системах.....	112
3.1.4. Линейные и инвариантные к сдвигу многомерные системы.....	112
3.2. Двумерные линейные фильтры.....	113
3.2.1. Оператор двумерного «скользящего среднего»	116
3.2.2. Лапласиан для «восьми соседей».....	116
3.2.3. Оператор выделения линий (контуров) в изображении	118
3.2.4. Оператор двойного дифференцирования.....	119
3.2.5. Оператор выделения малоразмерных объектов из шумов и фонов, подчиняющихся нормальному закону распределения	119
3.2.6. Градиентные операторы выделения перепадов яркости в различных пространственных направлениях	120
3.3. Рекурсивная обработка изображений.....	120
3.3.1. Интервальное интегрирование	121
3.3.2. Интервальное дифференцирование	122
3.3.3. Квазиоптимальная фильтрация малоразмерных объектов из шумов.....	123
3.3.4. Трапецеидальная ИХ.....	124
3.3.5. Двухкаскадный РСЦФ	124
3.4. Нелинейная обработка изображений.....	131
3.4.1. Ранговая обработка изображений	131
3.5. Двумерная децимация данных	137
4. СПЕЦИАЛИЗИРОВАННЫЕ УСТРОЙСТВА ДЛЯ ЦИФРОВОЙ ФИЛЬТРАЦИИ СИГНАЛОВ.....	144
4.1. Аппаратурное построение КИХ-фильтра прямой формы.....	144
4.2. Параллелизм при построении КИХ-фильтров прямой формы.....	147
4.3. Каскадная форма КИХ-фильтра.....	149
4.4. Прямая форма КИХ-фильтра с высоким уровнем параллелизма.....	152
4.5. Прямая форма построения БИХ-фильтров	154
4.6. Каскадная форма БИХ-фильтров.....	156
5. Цифровое сжатие видеосигналов.....	159
5.1. Пространственная и временная избыточность.....	159
5.2. Дискретное косинусное преобразование	159
5.3. Взвешивание	160
5.4. Кодирование с переменной длиной слова	161
5.5. Внутрикадровое кодирование	161
5.6. Межкадровое кодирование.....	162
5.7. Компенсация движения.....	163

5.8. I-видеокадры	165
5.9. Стандарты <i>MPEG</i>	166
5.9.1. Применения <i>MPEG</i>	166
5.9.2. Профили и уровни	167
5.9.3. <i>MPEG-1</i> и <i>MPEG-2</i>	169
5.9.4. Двухнаправленное кодирование	169
5.9.5. Типы данных	171
5.9.6. Структура цифрового потока <i>MPEG</i>	171
5.9.7. Системный уровень	173
ЛИТЕРАТУРА	175

ВВЕДЕНИЕ

Цифровая обработка сигналов (ЦОС) нашла широкое применение в следующих сферах деятельности: телевидении, радиолокации, связи, метеорологии, сейсмологии, медицине, анализе речи и телефонии, а также при обработке изображений и полей различной природы (фотографических, телевизионных, радиолокационных, рентгеновских, тепловых и т. д.). В некоторых сферах экономической деятельности, например, таких как банковское дело, обработка цифровых финансовых потоков имеет принципиальное значение.

Развитие вычислительной и микропроцессорной техники приводит к созданию все более надежного, быстродействующего, миниатюрного, качественного и недорогого оборудования. Цифровые технологии стали столь массовыми, что их используем в обыденной жизни особо не замечая: сотовый телефон, проигрыватель компакт-дисков, компьютер и т. д. В то же время для специалиста-радиоинженера эти устройства не могут оставаться «черными ящиками» и необходимо знать теорию ЦОС и уметь практически использовать ее при проектировании цифровых устройств обработки сигналов.

Многие годы ЦОС была «спрятана» в виде скромного раздела дисциплины «Радиотехнические цепи и сигналы», но в последнее десятилетие в ведущих ВУЗах России дисциплина «Цифровая обработка сигналов» сформировалась в виде самостоятельного курса. В ТУСУРе на радиотехническом факультете по инициативе кафедры телевидения и управления эта дисциплина преподается с 1992 г.

За 10 лет наряду с лекционным курсом разработаны: компьютерный лабораторный практикум по ЦОС, ряд практических и расчетных заданий, а также курсовых и дипломных проектных заданий. Введена в эксплуатацию лаборатория цифровых сигнальных процессоров (DSP) на оборудовании фирмы Texas Instruments, поставленного ТУСУРу по Университетской программе, проводимой этой фирмой. Таким образом, можно констатировать, что дисциплина «Цифровая обработка сигналов», в настоящее время, подкреплена учебно-научной базой и ее изучение может быть многоуровневым.

К ЦОС относят различные методы преобразования информации, такие как фильтрация (сглаживание) и предсказание, кодирование и сжатие, обнаружение и коррекция ошибок. Все большее количество методов обработки информации, применявшихся ранее в аналоговой технике, находят практическое воплощение в дискретном или цифровом виде.

Принято, что название определяет сущность предмета. Каждое из слов в аббревиатуре ЦОС является ключевым. **Сигнал** (лат. – *signum*) – знак, физический процесс (или явление), несущий сообщение (информацию о каком-либо событии, состоянии объекта наблюдения) либо передающий команды управления, указания, оповещения и т. д. **Обработка** – видоизменение, совокупность процессов (операций) при которых происходит формоизменение без нарушения содержания. **Цифра** (познелат. – *cifra*) – знак для обозначения

числа. Таким образом, ЦОС реализует ряд преобразований входного множества чисел в выходное множество, причем вид преобразования определяется конкретной задачей. Например, при фильтрации сигналов выполняются процедуры разделения (изменения) колебаний различных частот в спектре входного сигнала в желательном (для решения поставленной задачи) направлении.

Еще Ж. Маркос в своей книге «Дискретизация и квантование» обращал внимание на то, что через призму физики И. Ньютона окружающий нас мир кажется на редкость гладким. Если математики и вводили понятие разрывной функции, то для физиков (радиотехников) эта особенность не имела аналогов в окружающей природе. После появления работ А. Эйнштейна открылся мир квантовой физики – и с тех пор хотя мир и не изменился, но он становится для нас все более дискретным.

Фундаментальной основой для развития радиотехники и цифровой обработки сигналов, в том числе, стало появление теоремы Котельникова и теоремы Шеннона. Оказывается, чтобы изучить какое-то явление, не нужно знать его целиком, а можно лишь наблюдать его фрагменты (отсчеты) как во времени, так и в пространстве. Например, если мы желаем передать на расстояние сведения о непрерывном процессе (явлении) и результат восстановить на приемной стороне, то в ряде случаев передача не всего континуума данных, а лишь ограниченного множества чисел более предпочтительна. То есть в случае передачи определенного (достаточного) множества чисел на приемной стороне мы обеспечиваем более достоверное воспроизведение информации, чем при передаче всех сведений. В теории информации это объясняется тем, что все данные сверх некоего оптимума вносят шумы (искажают информацию), а не дополняют информацию о процессе (явлении). Таким образом, ЦОС по отношению к аналоговой обработке сигналов имеет принципиальное преимущество, которое усиливается еще и тем, что цифровая реализация устройств обработки стабильнее, точнее и более универсальна, а в настоящее время также конкурирует по стоимости с аналоговыми реализациями устройств обработки сигналов. В табл. В.1 сведены основные характеристики аналоговых фильтров (АФ) и цифровых фильтров (ЦФ).

Данное учебное пособие написано с использованием материалов отечественных и зарубежных авторов, а также методических и научных работ автора. В перечень литературы включен лишь ограниченный список [1-13], в который входят учебные пособия [1-3, 11, 12], монографии, ставшие библиографической редкостью [4-9] и специализированные публикации по разделам дисциплины ЦОС [10, 13].

Дисциплина ЦОС должна изучаться студентами после усвоения дисциплин: «Математика», «Радиотехнические цепи и сигналы» (РТЦиС) и «Цифровые устройства и микропроцессоры» (ЦУиМП). На рис. В.1 приведена структура дисциплины ЦОС и разделы других дисциплин, которые необходимо знать для ее успешного изучения. Именно этой структуре и подчинено изложение материала в данном учебном пособии.

Таблица В.1

Основные характеристики аналоговых и цифровых фильтров

№ п/п	Характеристики	АФ	ЦФ
1.	Переменные	Непрерывные по времени и континуальные уровню	Дискретные (квантованные) по времени, по уровню
2.	Математические операции	$d/dt, \int dt, \pm, \times, :$	Задержка на $T, 2T, \dots, MT; \pm, \times, :$
3.	Уравнения	Дифференциальные уравнения (линейные)	Разностные уравнения (линейные)
4.	Характерные отклики	Затухающие $\sin, \cos, \exp$ (устойчивые)	Выборки из затухающих $\sin, \cos, \exp$ (устойчивые)
5.	Суперпозиция, свертка	Для линейных фильтров соблюдается	Для линейных фильтров соблюдается
6.	Преобразования в частотной области	Фурье	Дискретное преобразование Фурье (ДПФ)
7.	Частотная область	На плоскости $p = j\omega$; $-\infty \leq \omega \leq \infty$.	На плоскости $z = e^{j\omega T}$; T – период дискретизации; $-\pi/T < \omega < \pi/T$
8.	Ограничения по большому сигналу	Нелинейность усилительных и преобразовательных устройств, их перегрузка	Переполнение арифметических устройств, нелинейность квантователя (АЦП)
9.	Ограничения по малому сигналу (шумы)	Тепловой и дробовой шумы, фоновые наводки	Квантование данных (шумы квантования) и результатов выполнения арифметических операций. Предельные циклы и «мертвые зоны»
10.	Стоимость	Установившаяся	Постоянно уменьшается
11.	Быстродействие	До «оптических» частот	Свыше 1 ГГц
12.	Технологичность	Исходные допуски, тепловой дрейф, старение элементов, нестабильность питания.	Округление коэффициентов цифровых фильтров (независимость от температуры, времени, дрейфа питающих напряжений, старения элементов)

СТРУКТУРА ДИСЦИПЛИНЫ ЦИФРОВОЙ ОБРАБОТКИ СИГНАЛОВ (ЦОС)

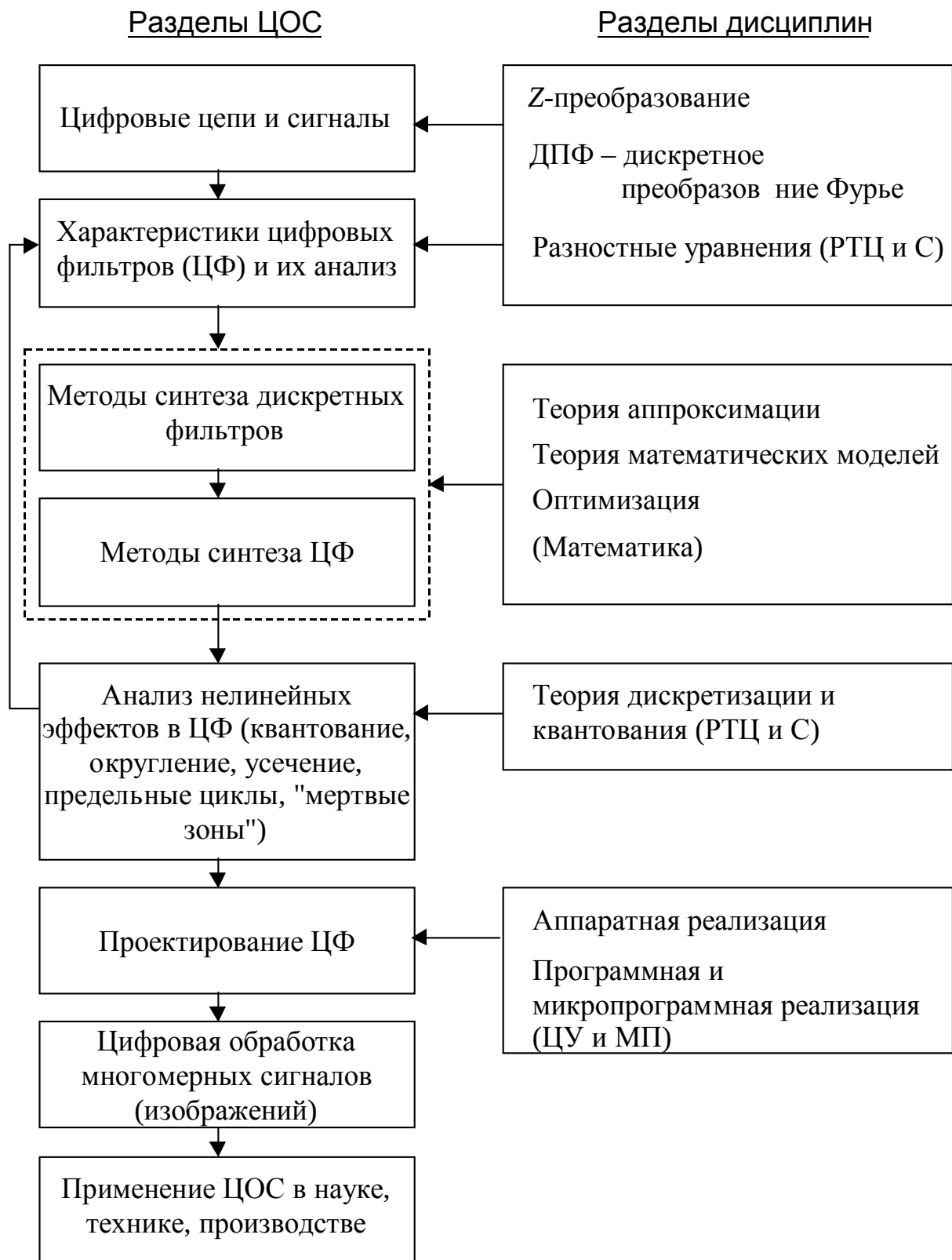


Рис. В.1.

Однако знания, полученные при изучении теоретической части дисциплины ЦОС, необходимо превратить в навыки и умения проектирования цифровых устройств обработки сигналов. Эту задачу следует решать читателю, как подсказывает философия о познании мира, используя индуктивный, традиционный и дедуктивный методы познания. Более прагматично выразился П.Л. Капица: «Все познание мира заключается в трех словах: теория, эксперимент, практика». Автор желает читателю пройти тот продуктивный путь, который проходит каждый, желающий понять мир и что-то в нем изменить.

В этой книге нет предисловия, где обычно выражают благодарность коллегам, поэтому автор перенес слова признательности во введение. Автор выражает благодарность сотрудникам кафедры телевидения и управления ТУСУР: доценту А.Г. Костевичу за полезные советы, критические замечания и дополнения, с ним автор в течение ряда лет обеспечивает проведение дисциплины «Цифровая обработка сигналов» для студентов радиотехнического факультета ТУСУР; аспиранту И.В. Гальчуку за помощь в оформлении и подготовке рукописи данного учебного пособия к изданию.

Обращаясь к читателю, хочется повторить фразу Р. Фейнмана, которую он вынес в эпилог своих знаменитых лекций: «Обучение редко приносит плоды кому-либо, кроме тех, кто предрасположен к нему, но им оно почти не нужно».

1. АНАЛИЗ ЦИФРОВЫХ ФИЛЬТРОВ

1.1. Цифровые цепи и сигналы

1.1.1. Цифровые сигналы

Рассмотрим одномерные цифровые сигналы, порожденные аналоговыми сигналами $x_a(t)$, $t_1 < t < t_2$, причем t_1 – часто принимают равным нулю, а t_2 – устремляют в бесконечность.

Аналоговый сигнал обычно ограничивают по амплитуде, т.е.

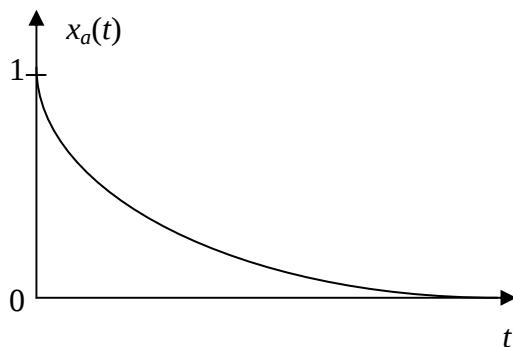


Рис. 1.1. – Аналоговый сигнал (экспонента)

$A_1 < x_a(t) < A_2$ – интервал изменения амплитуд, например:

изображенный на рис. 1.1 сигнал $x_a(t) = A_0 e^{-\alpha t}$, $\alpha > 0$, $A_0 = 1$, $t \geq 0$.

Из исходного аналогового сигнала можно выбрать бесконечное число выборок как по времени, так и по амплитуде.

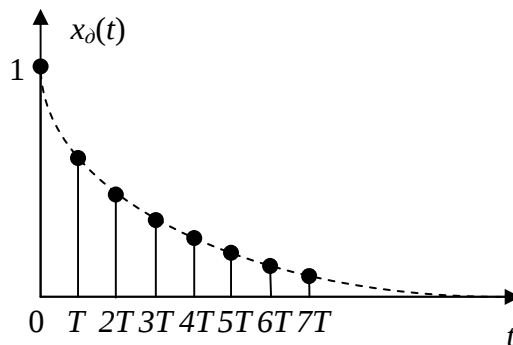


Рис. 1.2. – Дискретный сигнал (решетчатая функция)

Дискретные сигналы описываются решетчатыми функциями или последовательностями $x_d(nT)$ (рис. 1.2). В данном случае $x_d(nT) = A_0 e^{-\alpha \cdot nT}$, где $T = \text{const}$ – интервал дискретизации, $n = 0; 1; 2; \dots$, $f_d = \frac{1}{T}$ – частота дискретизации.

Сама функция может принимать произвольные значения на некотором интервале. Эти значения называются «выборками» или «отсчетами» решетчатой функции. Так, последовательность из трех отсчетов можно записать, например, в виде $x_d(0) = 1$, $x_d(T) = -2$, $x_d(2T) = 3$. Эту же последовательность можно также записать в виде $x_d(nT) = \{1; -2; 3\}$. Период дискретизации часто принимают равным $T = 1$ с. Для абстрактных последовательностей время T опускают (т. е. пишут $x_d(n)$).

Над решетчатыми функциями выполняют прямое Z-преобразование.

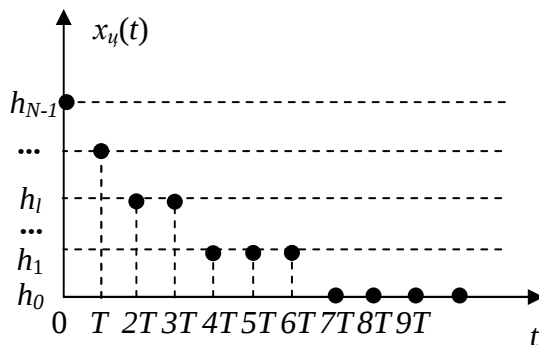


Рис. 1.3. – Цифровой сигнал (квантованная решетчатая функция)

Цифровые сигналы (ЦС) – это квантованные по уровню дискретные сигналы, которые описываются *кван-*

тованными решетчатыми функциями или квантованными последовательностями $x_q(nT)$, принимающими в дискретные моменты времени лишь ряд дискретных значений – уровней квантования $h_0, h_1, h_2, \dots, h_l, \dots, h_{N-1}$ (при этом уменьшается число выборок по амплитуде, чем больше N – тем выше точность, но увеличивается сложность цифрового устройства, т.к. требуется больше разрядов (рис. 1.3)).

Правильный выбор интервалов дискретизации по времени и по уровню (по амплитуде) очень важен при разработке систем ЦОС. Чем меньше интервал дискретизации T , тем точнее дискретизированный сигнал соответствует исходному непрерывному. Однако при уменьшении интервала дискретизации по времени возрастает число отсчетов и, для сохранения общего времени обработки сигнала неизменным, приходится увеличивать скорость обработки, что не всегда возможно. При уменьшении шага квантования по амплитуде требуется больше разрядов для представления сигнала, вследствие чего устройство становится более сложным и громоздким.

Связь между решетчатой функцией $x_o(nT)$ и квантованной функцией $x_q(nT)$ определяется нелинейной функцией квантования $x_q(nT) = F_k[x_o(nT)]$.

Существуют различные способы выбора функции квантования. В простейшем случае используется квантование с постоянным шагом $\Delta h = h_l - h_{l-1} = \text{const}$. Функция квантования имеет вид:

$$x_q(nT) = F_k[x(nT)] = \begin{cases} h_0 & \text{при } x(nT) \leq \frac{h_0 + h_1}{2}, \\ h_l & \text{при } \frac{h_{l-1} + h_l}{2} < x(nT) \leq \frac{h_l + h_{l+1}}{2}, \\ h_{N-1} & \text{при } x(nT) > \frac{h_{N-2} + h_{N-1}}{2}. \end{cases}$$

В дальнейшем мы оперируем не с самими отсчетами, а с их кодами. Каждый из уровней квантования кодируется числом. Обычно при кодировке используются двоичные числа (0 и 1), и квантованные отсчеты $x_q(nT)$ кодируются m -разрядными двоичными числами, например:

$$x_q(0) = 0001_2, x_q(T) = 0010_2, x_q(2T) = 0011_2 \text{ и т.д.}$$

Число уровней квантования N и наименьшее число разрядов m двоичных чисел, кодирующих эти уровни, связаны соотношением:

$$m = \text{Int}(\log_2 N),$$

где $\text{Int}(A)$ – наименьшее целое число, не меньшее, чем число A . Например, $\text{Int}(3,1) = 4$. Таким образом при $N = 4$ получим $m = 2$.

На рис. 1.4 приведена структурная схема обработки аналогового сигнала с использованием цифрового фильтра (ЦФ).

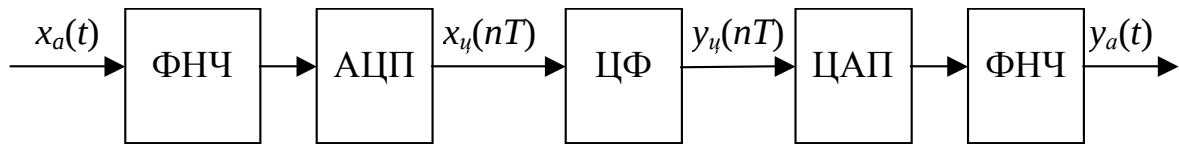


Рис. 1.4. – Структурная схема цифровой обработки аналогового сигнала

1.1.2. Аналого-цифровые и цифро-аналоговые преобразователи

Операции дискретизации по времени и квантования по уровню выполняются в аналого-цифровых преобразователях (АЦП). Структурную схему АЦП в самом общем виде можно представить как на рис. 1.5.

В качестве примера рассмотрим функциональную схему быстродействующего 6-разрядного АЦП параллельного типа (аналог К1107ПВ1) (рис. 1.6):

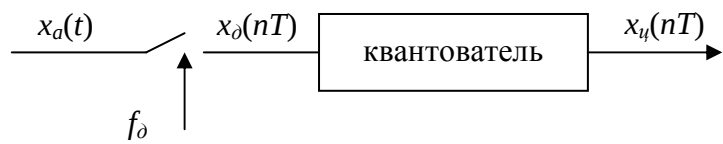


Рис. 1.5. – Структурная схема АЦП

$m = 6$ – число разрядов (К1107ПВ2 – 8 разрядов);

$N = 2^6$ – число уровней квантования (64 уровня);

f_∂ – частота дискретизации или тактовая частота (для К1107ПВ3

$f_\partial = 100$ МГц).

АЦП серии К1107 применяются в телевидении, радиолокации, антенных системах, в устройствах ввода-вывода информации и др.

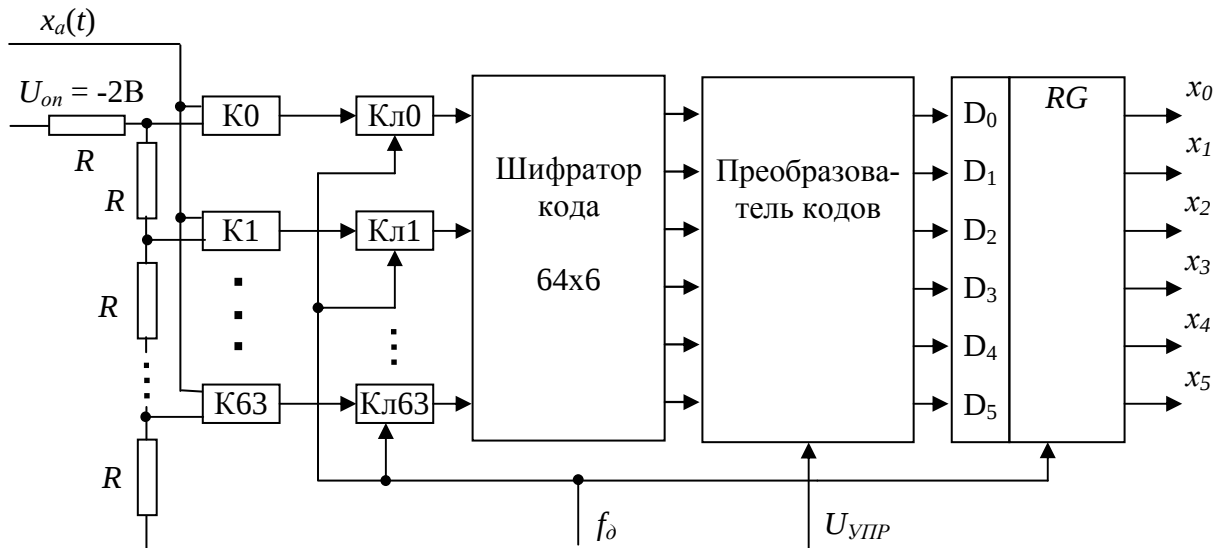


Рис. 1.6. – Функциональная схема АЦП (аналог К1107ПВ1)

В АЦП сигнал проходит несколько этапов преобразования:

- квантование по уровню;
- дискретизация по времени;
- кодирование (представление двоичным кодом);
- модификация кода (прямой код, используемый при умножении/делении, и обратный код, используемый при сложении/вычитании).

Основные параметры АЦП:

1. Число разрядов выходного кода определяет количество уровней квантования. Для m -разрядного двоичного кода это количество составляет 2^m .
2. Разрешающая способность. Определяется минимальным входным напряжением, соответствующим изменению выходного кода на единицу младшего разряда (МР). Характеризует порог чувствительности, при котором происходит смена выходного кода. Для m -разрядного двоичного кода при максимальном входном напряжении $U_{вх\ max}$ разрешающая способность равна $U_{вх\ max}/(2^m - 1)$.
3. Нелинейность Δ_n , т.е. максимальное отклонение реальной передаточной характеристики АЦП от идеальной (линейной). Выражается в процентах или долях МР.
4. Абсолютная погрешность преобразования $\Delta_{ми}$ (погрешность усиления и смещения нуля).
5. Время преобразования, т.е. интервал времени от момента заданного изменения сигнала на входе до появления на выходе установившегося кода (часто используют термин – частота преобразования).

Цифро-аналоговые преобразователи (ЦАП) обеспечивают преобразование входной информации в цифровой форме (двоичном коде) в выходную информацию в аналоговой форме. Входной код может быть последовательным или параллельным. Последний случай более распространен на практике.

Принцип действия ЦАП состоит в суммировании эталонных токов или напряжений, т.е. каждому уровню цифрового сигнала соответствует определенный набор эталонных токов (напряжений), которые в сумме представляют собой значение аналогового сигнала.

В большинстве серийных отечественных преобразователей осуществляется преобразование входного кода сначала в ток, а затем в напряжение. Обобщенная структурная схема ЦАП, соответствующая таким преобразователям, приведена на рис. 1.7. В нее входят следующие основные узлы: резистивная матрица, с помощью которой формируются токи, соответствующие входному коду ЦАП; токовые ключи, подключающие в соответствии с входным кодом цепи резистивной матрицы; устройство для согласования входа ЦАП с цифровыми микросхемами (ТТЛ, ЭСЛ, КМДП, микропроцессорами и др.). В ряде ЦАП используется дополнительный узел – генераторы тока, обеспечивающие точное поддержание токов, соответствующих входному коду. Кроме того, в структурную схему входит источник опорного напряжения.

Все перечисленные узлы обеспечивают преобразование кода в ток. Для обеспечения выполнения функции преобразователя код–напряжение необходим еще операционный усилитель (ОУ), преобразующий ток в напряжение. ОУ и источник опорного напряжения сложно выполнить технически на одном кристалле с остальной частью преобразователя, поэтому эти узлы часто выполняют в виде отдельных микросхем.

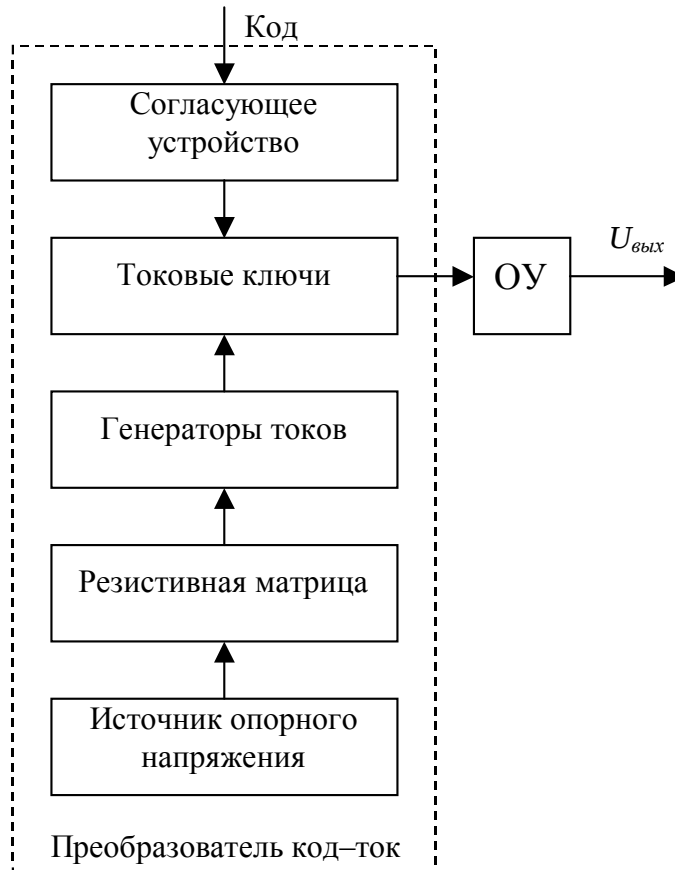


Рис. 1.7. – Структурная схема ЦАП

В суммировании участвуют только те токи или напряжения, для которых в соответствующих разрядах двоичного кода стоит 1. Для получения эталонных токов используют прецизионные матрицы $R-2R-4R-8R-\dots$ либо $R-2R$ (см. рис. 1.8).

На рис. 1.8 а:

$U_{\text{вых}} = KU_{\text{он}}(a_1 2^{-1} + a_2 2^{-2} + \dots + a_m 2^{-m})$, где $U_{\text{он}}$ – опорное (эталонное) напряжение;

$a_1, a_2, \dots, a_m$ – двоичные разряды, принимающие значения 0 или 1;

K – коэффициент пропорциональности $\left(K = -\frac{2R_{\text{OC}}}{R} \right)$;

R_{OC} – сопротивление обратной связи (ОС) операционного усилителя;

$I_0 = \frac{2U_{\text{он}}}{R} \cdot \frac{2^m - 1}{2^m}$ – опорный ток, при $m \rightarrow \infty$ $I_0 \rightarrow \frac{2U_{\text{он}}}{R}$.

На рис. 1.8 б: $I_0 = \frac{U_{\text{он}}}{R}$; $I_{\Sigma} = I_0 \sum_{i=1}^m I_i = I_0 \sum_{i=1}^m a_i 2^{-i}$ – опорный и выходной

токи;

$U_{\text{вых}} = -R_{\text{OC}} I_{\Sigma} = -U_{\text{он}} \frac{R_{\text{OC}}}{R} \sum_{i=1}^m a_i 2^{-i}$ – выходное напряжение.

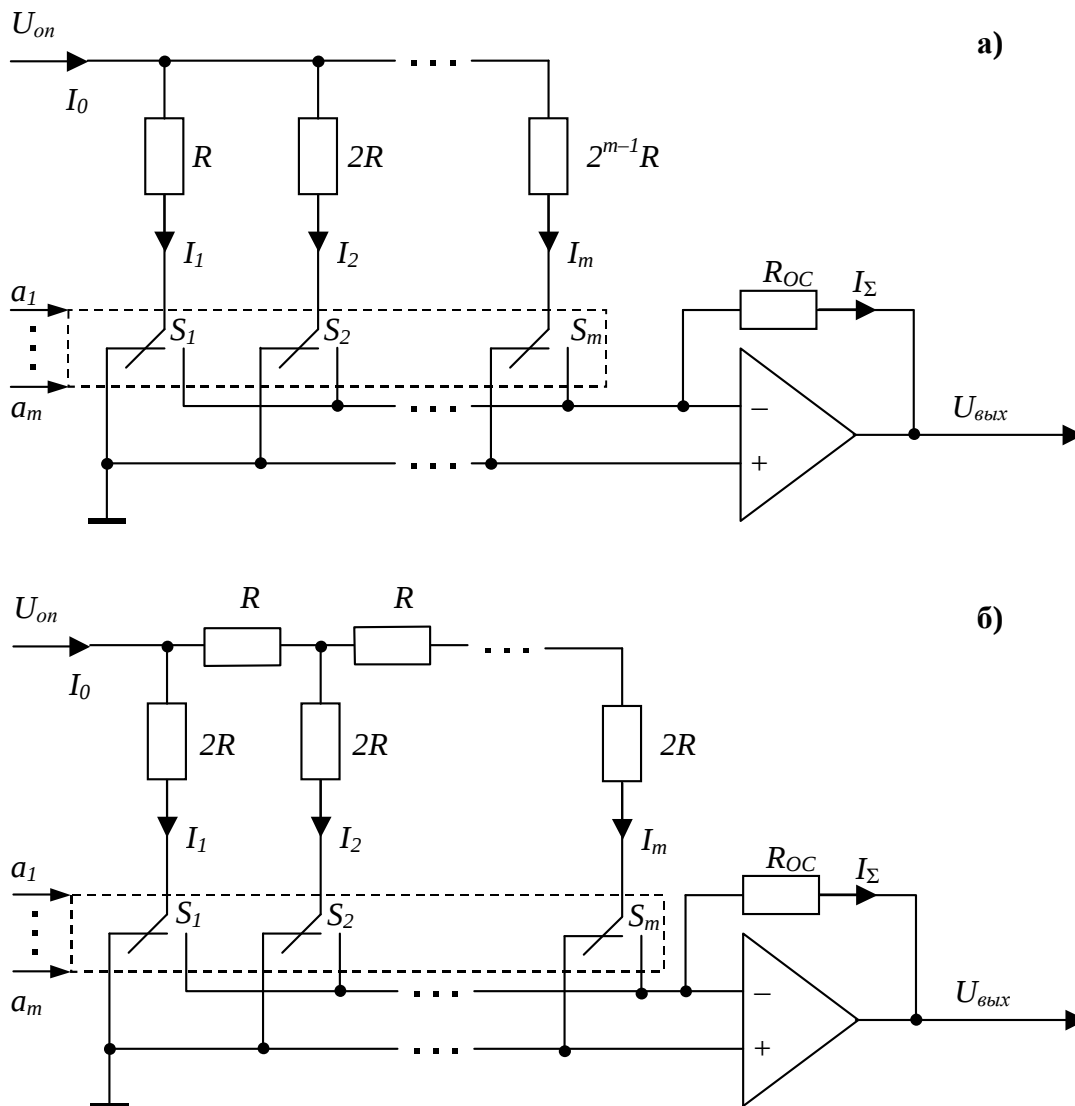


Рис. 1.8. – Умножающий ЦАП (аналог К572ПА1)

При $m = 10$ получим $N = 2^{10} = 1024$ уровней (разрядов) выходного сигнала. Время установления (см. ниже) $t_{ycm} = 5$ мкс.

Обычно ЦАП гораздо дешевле АЦП и их используют при построении АЦП. Один из вариантов построения АЦП с использованием ЦАП приведен на рис. 1.9. Здесь ГТИ – генератор тактовой частоты; К – компаратор; ФС – формирователь сигнала готовности.

При $U_{зан} < x(t)$ на выходе компаратора логическая единица, в противном случае – логический 0.

По спаду (заднему фронту) сигнала на выходе компаратора формируется сигнал «готовность» и происходит запись двоичного кода в выходной регистр.

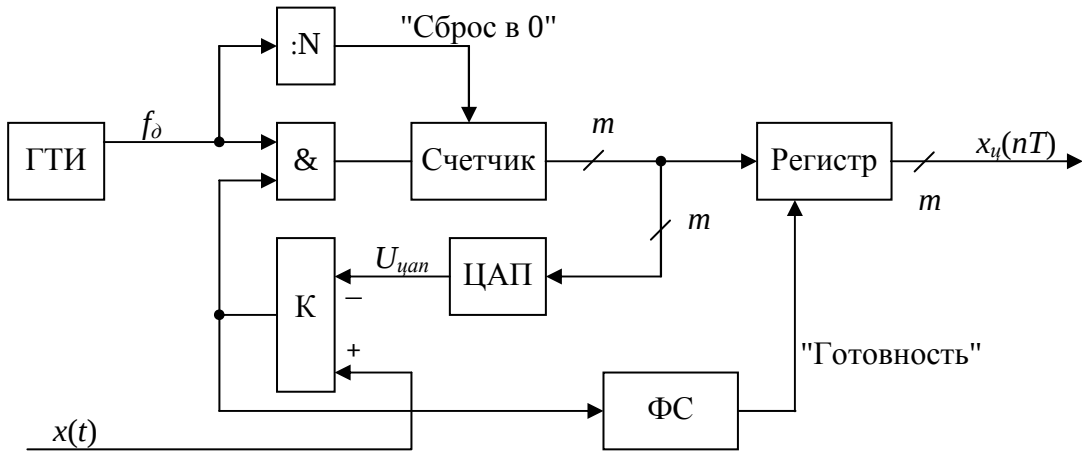


Рис. 1.9. – АЦП последовательного действия с обратной связью через ЦАП (аналог К572ПВ1)

Основные параметры ЦАП:

1. Число разрядов управляющего входного кода – m ; $N = 2^m$ – число уровней выходного сигнала.

2. Разрешающая способность – характеризует возможность ЦАП различать смежные значения входного кода – 1 МР. Разрешающая способность имеет величину $\frac{U_{on}}{2^m - 1}$ и определяется в % от полной шкалы.

3. Абсолютная погрешность преобразования – $\Delta_{\text{ми}}$. Это отклонение выходного напряжения от расчетного значения в конечной точке характеристики преобразования. Включает в себя – смещение нуля, изменение коэффициента передачи «код-напряжение» и др. Выражается в долях МР или в % от полной шкалы. Типичное значение погрешности ЦАП д.б. меньше половины МР.

4. Нелинейность (Δ_L) – максимальное отклонение реальной характеристики преобразования от теоретической характеристики (прямой линии, соединяющей точку нуля и максимум выходного сигнала). Подстройкой не компенсируется.

5. Дифференциальная нелинейность ($\Delta_{\text{ДН}}$) – максимальное отклонение разности двух значений аналогового сигнала, соответствующих соседним кодам, от значения МР. Не компенсируется подстройкой. Определяется долями МР или в % от полной шкалы.

6. Время установления (преобразования) – $t_{уст}$ – интервал времени от подачи входного кода до момента, когда выходной сигнал достигнет установившегося значения с заданной погрешностью (обычно половина МР). Это время определяет общее быстродействие ЦАП.

K572ПA1: $m = 10$, $\Delta_I = \pm 0,1 \%$, $t_{vcm} = 5$ мкс, $f_{np} = 200$ кГц;

К1118ПАЗ: $m = 8$, $\Delta_{II} = \pm 0,195 \%$, $t_{ycm} = 0,01$ мкс, $f_{np} = 100$ МГц.

Принятые обозначения сигналов и характеристик АЦП и ЦАП:

b (b)	– число разрядов (на рисунках это m);
$\Delta_{\text{пш}} (\Delta_{\text{FS}})$	– абсолютная погрешность;
$\Delta_L (\Delta_L)$	– нелинейность;
$\Delta_{\text{лд}} (\Delta_{\text{LD}})$	– дифференциальная нелинейность;
$t_{\text{уст}} (t_{\text{S}})$	– время установления;
$U'_{\text{вх}} (U_{1\text{H}})$	– входное напряжение высокого уровня (логическая 1);
$U^0_{\text{вх}} (U_{1\text{L}})$	– входное напряжение низкого уровня (логический 0);
$U_{\text{п}} (U_{\text{сс}})$	– напряжение питания;
$I_{\text{ном}} (I_{\text{сс}})$	– потребляемый ток;
$U_{\text{оп}} (U_{\text{REF}})$	– опорное напряжение;
$U_{\text{вых}} (U_{\text{ORN}})$	– выходное напряжение;
$I_{\text{вых}} (I_{\text{ORN}})$	– выходной ток.

1.1.3. Цифровые системы обработки сигналов

Цифровыми системами обработки сигналов называют системы, выполняющие однозначное преобразование входного массива данных $x_u(nT)$ в выходной массив данных (рис. 1.10).

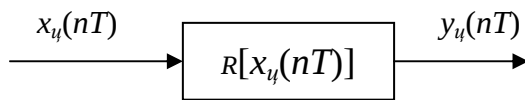


Рис. 1.10. – Цифровая система обработки сигналов

Здесь $R[\cdot]$ – оператор преобразования;

$$y_u(nT) = R[x_u(nT)];$$

$y_u(nT)$ – отклик системы или реакция системы на входное воздействие.

Такое определение описывает очень широкий класс систем. Мы же чаще будем рассматривать более узкий класс систем: стационарных, линейных, физически реализуемых и устойчивых. Дадим определение таких систем.

Цифровая система называется стационарной (инвариантной во времени), если ее параметры не изменяются во времени. Это означает, что вид отклика на заданное входное воздействие не зависит от времени его прихода на вход системы. Пусть $x(nT) \equiv y(nT) = 0$ при $n < 0$ – невозбужденная система или система с нулевыми начальными условиями, тогда, например, $y(nT) = R[x(nT)] = x(nT - T) + 0,5x(nT - 2T)$ – оператор преобразования справедлив при любом значении n .

Нестационарная система

$$y(nT) = R[x(nT)] = 2nx(nT).$$

Здесь $y(nT) = R[x(nT - kT)] = 2(n - k) \cdot x(nT - kT) \neq y(nT - kT)$.

Линейной цифровой системой называется система, для которой выполняется принцип суперпозиции. Пусть, например, $\alpha \cdot x_1(nT) + \beta \cdot x_2(nT)$ – сложное входное воздействие, а операторы преобразования определены следующим образом: $R[x_1(nT)] = y_1(nT)$ и $R[x_2(nT)] = y_2(nT)$.

Тогда реакция такой системы на сложное воздействие будет равна взвешенной с теми же коэффициентами α и β сумме реакций на отдельные воздействия

$$\begin{aligned} y(nT) &= R[\alpha \cdot x_1(nT) + \beta \cdot x_2(nT)] = \alpha \cdot R[x_1(nT)] + \beta \cdot R[x_2(nT)] = \\ &= \alpha \cdot y_1(nT) + \beta \cdot y_2(nT) \end{aligned}$$

при всех ненулевых α и β и всех доступных для данной системы воздействий $x_1(nT)$ и $x_2(nT)$.

Пример: $y(nT) = R[x(nT)] = 0,5x(nT - T) + 0,1x(nT - 2T)$. Пусть сложное входное воздействие $x(nT) = \alpha x_1(nT) + \beta x_2(nT)$, тогда выходной сигнал представляет 2 сигнала:

$$\begin{aligned} y(nT) &= 0,5[\alpha x_1(nT - T) + \beta x_2(nT - T)] + 0,1[\alpha x_1(nT - 2T) + \beta x_2(nT - 2T)] = \\ &= \alpha[0,5x_1(nT - T) + 0,1x_1(nT - 2T)] + \beta[0,5x_2(nT - T) + 0,1x_2(nT - 2T)] = \\ &= \alpha y_1(nT) + \beta y_2(nT). \end{aligned}$$

Нелинейная система – это система, операторы преобразования которой представляют собой математически нелинейные функции, например,

$$y(nT) = x^2(nT) - \text{нелинейная, но стационарная система.}$$

Физически реализуемыми системами называются системы, у которых текущий отсчет выходного сигнала $y(nT)$ формируется из предыдущих отсчетов и текущего отсчета входного сигнала и предыдущих отсчетов выходного сигнала. Такая система называется *каузальной*, т.е. сохраняющей причинно-следственные связи. Например, $y(nT) = 0,5x(nT - 2T)$ – физически реализуемая система, а $y(nT) = 0,5x(nT + 2T)$ – физически нереализуемая система. Иначе можно сказать, что *физически реализуемыми* системами называются системы, реакция которых в данный момент времени не зависит от значений воздействий в последующие моменты времени.

Устойчивыми называют системы, выходные отклики которых ограничены при любом ограниченном входном воздействии, т.е. $\sum_{n=0}^{\infty} |h(nT)| < \infty$, где

$h(nT)$ есть импульсная характеристика. Или по другому, система устойчива, если все полюсы системной функции $H(z)$ расположены внутри единичного круга z -плоскости. Понятие устойчивости требует уточнения. Например, *интегратор – условно устойчивая система*, т.к. ограничение на амплитуду выходного воздействия требует ограничения действия входного сигнала по времени.

В дальнейшем устройства, реализующие ЦОС, мы будем называть цифровыми фильтрами (ЦФ).

Цифровым фильтром, как правило, называют линейную, инерционную, цифровую систему, которую можно использовать для фильтрации дискретных во времени и квантованных по уровню сигналов.

Как и другие системы обработки сигналов, цифровые фильтры могут быть стационарными или нестационарными, линейными или нелинейными, физически реализуемыми или нереализуемыми.

В табл. 1.1 приведены базовые операции, используемые для построения линейных ЦФ, а в табл. 1.2 – базовые сигналы, используемые в качестве входных воздействий.

Таблица 1.1

Базовые операции в линейных цифровых фильтрах

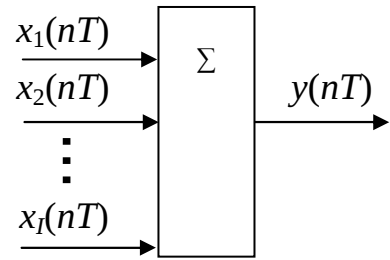
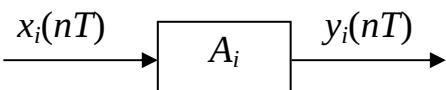
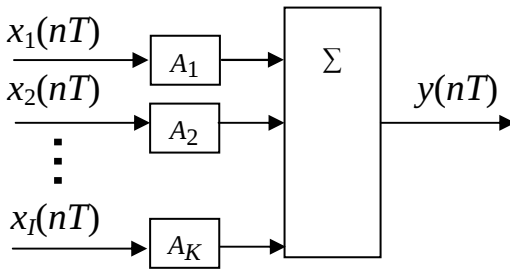
№ п/п	Операция	Математическая запись	Обозначения
1.	Суммирование, вычитание	$y(nT) = \sum_{i=1}^I \text{sgn}_i \cdot x_i(nT)$ $\text{sgn} = 1$ при суммировании $\text{sgn} = -1$ при вычитании	
2.	Умножение, деление на коэффициент	$y_i(nT) = A_i \cdot x_i(nT)$	
3.	«Полная» арифметика	$y(nT) = \sum_{i=1}^K A_i \cdot x_i(nT)$	
4.	Задержка сигнала на один либо k периодов дискретизации	$y(nT) = x(nT - T)$ $y(nT) = x(nT - kT)$	

Таблица 1.2

Базовые сигналы в цифровых фильтрах

№ п/п	Функция	Определение	График
1.	Единичный импульс (ЕИ)	$\delta(nT) = \begin{cases} 1, & n = 0 \\ 0, & n \neq 0 \end{cases}$	
1.1	Задержанный и масштабированный ЕИ	$A\delta(nT - kT) = \begin{cases} A, & n = k \\ 0, & n \neq k \end{cases}$	
2.	Единичная ступенчатая функция (можно сдвигать и масштабировать)	$u(nT) = \begin{cases} 1, & n \geq 0 \\ 0, & n < 0 \end{cases}$ $\delta(nT) = u(nT) - u(nT - T)$	
3.	Единичная наклонная функция	$r(nT) = \begin{cases} nT, & n \geq 0 \\ 0, & n < 0 \end{cases}$	

Аналитическая запись последовательности $x(nT)$ (фильтрующее свойство

ЕИ): $x(nT) = \sum_{k=0}^{\infty} x(kT) \cdot \delta(nT - kT)$ при $k < 0$, $x(kT) = 0$.

Пусть определена последовательность из трех отсчетов $x(nT) = \{1, -2, 3\}$. Тогда аналитическая запись последовательности может быть представлена в виде $x(nT) = \delta(nT) - 2\delta(nT - T) + 3\delta(nT - 2T)$.

Следует отметить, что аналитическая запись используется при вычислении прямого Z-преобразования, свертки, откликов и т.д. [4].

1.2. Дискретное преобразование Фурье (ДПФ)

Основные формулы для ДПФ:

1. Для непрерывного сигнала:

$$X_a(\omega) = \int_{-\infty}^{\infty} x_a(t) e^{-j\omega t} dt - \text{спектр непрерывного сигнала } x_a(t),$$

$$x_a(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} X_a(\omega) e^{j\omega t} d\omega - \text{обратное преобразование Фурье.}$$

2. Для дискретного (периодического) сигнала:

$$X(e^{j\omega T}) = \sum_{n=-\infty}^{\infty} x(nT) e^{j\omega nT} - \text{спектр дискретизированного сигнала,}$$

здесь: $\omega \in \left[-\frac{\omega_\delta}{2}; \frac{\omega_\delta}{2} \right]$ - период повторения спектра, при этом

$$\omega_\delta = \frac{2\pi}{T} - \text{угловая частота дискретизации;}$$

$$x(nT) = \frac{T}{2\pi} \int_{-\frac{\pi}{T}}^{\frac{\pi}{T}} X(e^{j\omega T}) e^{j\omega nT} d\omega - \text{обратное преобразование Фурье.}$$

Из теории дискретизации непрерывных сигналов следует, что реальные сигналы может быть описаны выборками как в спектральной, так и во временной областях. Непосредственно переход от дискретного сигнала к дискретному спектру и наоборот возможен с использованием дискретного преобразования Фурье.

Число отсчетов сигнала или спектра (т.е. число степеней свободы сигнала) обычно ограничивают величиной $N = 2^m$, где $m = 1, 2, 3, \dots$. Если $x(nT)$ вещественная функция, то в силу четности ДПФ число отсчетов спектра сокращается в 2 раза (одностороннее преобразование).

Одно из основных применений ДПФ это вычисление сигнала на выходе фильтра с заданной амплитудно-частотной характеристикой (АЧХ). При этом отсутствуют какие-либо ограничения на вид АЧХ. Она м.б., например, идеально прямоугольной, что невозможно реализовать с помощью обычных фильтров.

Использование ДПФ:

а) с помощью прямого ДПФ находим спектр $X(e^{j\omega T})$ входного сигнала $x(nT)$;

б) находим спектр выходного сигнала $Y(e^{j\omega T}) = X(e^{j\omega T}) \cdot H(e^{j\omega T})$, где $H(e^{j\omega T})$ – частотная характеристика фильтра;

в) с помощью обратного ДПФ находим выходной сигнал $y(nT)$.

Обработку сигналов с помощью ДПФ нельзя назвать цифровой фильтрацией в полном смысле слова. Обычные ЦФ, работающие в реальном времени, обрабатывают сигнал непрерывно по мере его поступления. А обработка ДПФ производится лишь после того, как будет полностью известен входной сигнал. Поэтому при использовании ДПФ выходной сигнал м.б. получен с некоторым запаздыванием по отношению к входному сигналу. Однако в ряде практических применений такое запаздывание не играет существенной роли. Если входной сигнал имеет большую длительность, то его обработку с помощью ДПФ можно производить по частям. Для повышения точности обработки обрабатываемые серии отсчетов сигнала могут частично перекрываться.

Непосредственное вычисление ДПФ требует выполнения большого объема вычислений – так для N отсчетов сигнала необходимо выполнить N^2 умножений и $N(N-1)$ сложений. Поэтому для больших N ($N > 1000$) целесообразнее применять алгоритмы быстрого преобразования Фурье (БПФ).

Так, например, согласно одному из алгоритмов сумму (см. выше уравнение для спектра дискретного сигнала) разбивают на две части – для четных и нечетных значений n . Затем в каждую часть вводят новые индексы для нумерации отсчетов и, поступая аналогично по отношению к новым индексам, разбивают каждую часть на две и т.д., пока не останутся части, содержащие по одному отсчету сигнала. Такой алгоритм называют «алгоритмом с прореживанием по времени».

Согласно еще одному алгоритму сумму в уравнении для спектра дискретного сигнала разбивают на части аналогичным способом, но не по четности индексов, а просто – пополам. Такой алгоритм называют «алгоритмом с прореживанием по частоте».

1.3. Z-преобразование

В задачах анализа и синтеза линейных дискретных систем широко применяются методы Z-преобразования.

Запишем сначала преобразование Лапласа для непрерывных сигналов:

$$F(p) = \int_0^{\infty} f(t) \cdot e^{-pt} dt, \quad t \geq 0.$$

Преобразование Лапласа можно применять практически для всех сигналов, поскольку все реально существующие сигналы имеют начало.

Применим преобразование Лапласа к дискретному сигналу, записанному в виде последовательности $\tilde{\delta}$ -функций:

$$f(t) = \sum_{n=0}^{\infty} f(nT) \cdot \tilde{\delta}(t - nT).$$

Подставляя выражение для $f(t)$ в выражение для $F(p)$ и меняя местами суммирование и интегрирование, получим дискретное преобразование Лапласа:

$$F(p) = \sum_{n=0}^{\infty} f(nT) \cdot e^{-pnT}.$$

Если положить $e^{pT} = z$, то в результате такой замены придем к *прямому Z-преобразованию*, которое обычно применяют для анализа дискретных сигналов, т.е.

$$F(z) = Z\{f(nT)\} = \sum_{n=0}^{\infty} f(nT) \cdot z^{-n}, \text{ где } n = 0, 1, 2, \dots$$

Это одностороннее Z-преобразование, где множество значений комплексной переменной $z = re^{j\varphi} = \alpha + j\beta$, при которых ряд $F(z)$ сходится, называется областью сходимости. Иначе говоря, область сходимости ряда определяется радиусом круга r в z -плоскости, вне которого ряд сходится. Для равномерной сходимости ряда достаточно, чтобы выполнялось условие:

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

Обратное преобразование $Z^{-1}\{\cdot\}$ позволяет определить значения дискретного сигнала по виду функции $F(z)$, т.е. можно сказать, что обратное Z-преобразование ставит в соответствие функции комплексной переменной $F(z)$ решетчатую функцию (последовательность) $f(nT) = Z^{-1}\{F(z)\}$, определяемую по формуле

$$f(nT) = \frac{1}{2\pi j} \oint F(z) \cdot z^{n-1} dz, \text{ где } \oint dz - \text{интеграл по контуру, расположен-}$$

ному в области сходимости $F(z)z^{n-1}$ и охватывающему начало координат в z -плоскости. Этот интеграл удобно вычислять при помощи теоремы о вычетах: функция $f(nT)$ определяется суммой вычетов подынтегральной функции в полюсах, расположенных в области, охватываемой контуром интеграла, т.е.

$$f(nT) = \sum_{k=1}^K \operatorname{Res}_{z=z_k} \{F(z)z^{n-1}\}, \text{ причем вычет в простом полюсе } z = z_k \text{ равен}$$

$$\operatorname{Res}_{z=z_k} \{F(z)z^{n-1}\} = \lim_{z \rightarrow z_k} \{(z - z_k)F(z)z^{n-1}\}.$$

Удобный способ вычисления обратного Z-преобразования заключается в разложении $F(z)$ на простые дроби (если $F(z)$ – рациональная функция):

$$F(z) = \sum_{s=1}^N \frac{\beta_s}{1 - \alpha_s z^{-1}}.$$

В этом случае, используя свойство линейности, найдем, что

$$f(nT) = \sum_{s=1}^N \beta_s (\alpha_s)^n.$$

Наиболее важными свойствами Z-преобразования являются линейность, сдвиг (задержка) последовательности, свертка последовательностей, перемножение последовательностей и выполнение равенства Парсеваля.

Линейность

Если последовательности $f_1(nT)$ и $f_2(nT)$ имеют, соответственно, Z-преобразования $F_1(z)$ и $F_2(z)$ и a_1, a_2 – постоянные, не зависящие от k коэффициенты, то $F(nT) = a_1 f_1(nT) + a_2 f_2(nT)$ имеет изображение

$$F(z) = a_1 F_1(z) + a_2 F_2(z).$$

Z-преобразование задержанной последовательности

Задержим дискретный сигнал $x(nT)$ во времени на величину интервала дискретизации T , т. е. $y(nT) = x(nT - T)$ – новый сигнал. Пусть также известно

Z-преобразование $x(nT)$: $X(z) = \sum_{n=0}^{\infty} x(nT) \cdot z^{-n}$. Тогда

$$Y(z) = \sum_{n=0}^{\infty} y(nT) \cdot z^{-n} = \sum_{n=1}^{\infty} x(nT - T) \cdot z^{-n} = \sum_{m=0}^{\infty} x(mT) \cdot z^{-(m+1)} = z^{-1} \cdot X(z).$$

Таким образом, запаздывание дискретного сигнала на величину T соответствует умножению Z-преобразования на z^{-1} и т.д.

Свертка последовательностей

Пусть $X_1(z) = Z\{x_1(nT)\}$, $X_2(z) = Z\{x_2(nT)\}$. Свертка последовательностей $x_1(nT)$ и $x_2(nT)$

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

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

Перемножение последовательностей

Если $X_1(z) = Z\{x_1(nT)\}$ и $X_2(z) = Z\{x_2(nT)\}$, то последовательность $y(nT) = x_1(nT) x_2(nT)$ имеет Z-преобразование $Y(z) = \frac{1}{2\pi j} \oint_C X_1(v) X_2\left(\frac{z}{v}\right) \frac{dv}{v}$, где контур C лежит внутри пересекающихся областей сходимости $X_1(v)$ и $X_2(z/v)$. Следствием этой формулы является преобразование Фурье (спектр) сигнала

$y(nT)$, т.е. $Y(e^{j\omega T}) = \frac{T}{2\pi} \int_{-\pi/T}^{\pi/T} X_1(e^{j\Psi T}) X_2(e^{j\omega T}) d\Psi$, где $X_1(e^{j\omega T})$ и $X_2(e^{j\omega T})$ – спектры сигналов $x_1(nT)$ и $x_2(nT)$.

Равенство Парсеваля

Если положить $y(nT) = x_1(nT)x_2^*(nT)$ (звездочка означает комплексно-сопряженную величину), то можно записать

$$\sum_{n=0}^{\infty} y(nT) = \sum_{n=0}^{\infty} x_1(nT)x_2^*(nT) = \frac{T}{2\pi} \int_{-\pi/T}^{\pi/T} X_1(e^{j\omega T}) X_2^*(e^{j\omega T}) d\omega.$$

При $x_1(nT) = x_2(nT) = x(nT)$ получим равенство Парсеваля

$$\sum_{n=0}^{\infty} |x(nT)|^2 = \frac{T}{2\pi} \int_{-\pi/T}^{\pi/T} |X(e^{j\omega T})|^2 d\omega.$$

В табл. 1.3 даны примеры прямого и обратного Z-преобразований.

1.4. Алгоритмы функционирования и формы реализации линейных цифровых фильтров

Линейный цифровой фильтр (ЛЦФ) – устройство, в котором выходные отсчеты сигнала представлены в виде линейной комбинации предыдущих отсчетов входного и выходного сигналов и текущего отсчета входного сигнала.

Основное разностное уравнение для ЛЦФ имеет вид:

$$y(nT) = \sum_{i=0}^M a_i \cdot x(nT - iT) + \sum_{j=1}^N b_j \cdot y(nT - jT), \quad n \geq 0.$$

Здесь T – период (интервал) дискретизации;

n – текущий номер отсчета сигнала, включая нулевой отсчет ($n = 0, 1, \dots$);

j – начинается с 1, чтобы уравнение было приведенным.

Первая сумма описывает нерекурсивную часть фильтра, вторая – рекурсивную часть фильтра.

Если a_i и $b_j \equiv \text{const}$, то ЛЦФ с постоянными параметрами, иначе ЛЦФ с переменными параметрами.

Большее из чисел M и N – является порядком фильтра.

Если $b_j \equiv 0$, то получаем нерекурсивные фильтры или фильтры с конечной импульсной характеристикой, так называемые КИХ-фильтры.

Фильтры без обратных связей, т.е. $b_j \equiv 0$ называют также однородными, трансверсальными или фильтрами с многоотводной линией задержки [1]. Если $b_j \neq 0$, то получаем класс рекурсивных фильтров.

Таблица 1.3

Z-преобразование широко используемых последовательностей

№ п/п	Функция	Z-образ
1.	$f(nT)$	$F(z)$
2.	$\delta(nT)$	1
3.	$u(nT)$	$\frac{z}{z-1}$
4.	$u(nT-T) \cdot K$	$\frac{K}{z-1}$
5.	$u(nT) \cdot K \cdot A^n$	$\frac{K \cdot z}{z-A}$
6.	$u(nT-T) \cdot K \cdot A^{n-1}$	$\frac{K}{z-A}$
7.	$u(nT) \cdot e^{-anT}$	$\frac{z}{z-e^{-aT}}$
8.	$r(nT)$	$\frac{T \cdot z}{(z-1)^2}$
9.	$u(nT) \cdot \sin(\omega nT)$	$\frac{z \cdot \sin(\omega T)}{z^2 - 2z \cdot \cos(\omega T) + 1}$
10.	$u(nT) \cdot \cos(\omega nT)$	$\frac{z \cdot (z - \cos(\omega T))}{z^2 - 2z \cdot \cos(\omega T) + 1}$
11.	$u(nT) \cdot e^{-anT} \cdot \sin(\omega nT)$	$\frac{z \cdot e^{-aT} \sin(\omega T)}{z^2 - 2z \cdot e^{-aT} \cos(\omega T) + e^{-2aT}}$
12.	$u(nT) \cdot e^{-anT} \cdot \cos(\omega nT)$	$\frac{z \cdot (z - e^{-aT} \cos(\omega T))}{z^2 - 2z \cdot e^{-aT} \cos(\omega T) + e^{-2aT}}$

На рис. 1.11 проиллюстрирован принцип формирования выходного сигнала в ЛЦФ.

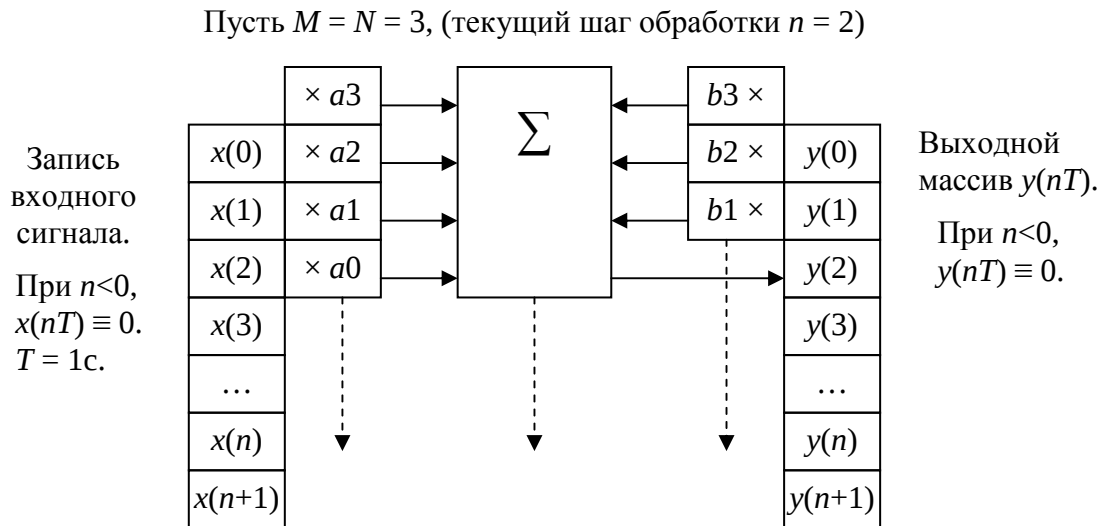


Рис. 1.11. – Формирование выходного сигнала в ЛЦФ

Данное разностное уравнение точно описывает алгоритмы функционирования не цифровых, а дискретных фильтров, т.е. не учитываются эффекты квантования. В то же время это разностное уравнение описывает все типы линейных дискретных фильтров (полосовые, режекторные, ФНЧ, ФВЧ и др.).

Формы реализации (структурные схемы) ЛЦФ м. б. различными: прямая (рис. 1.12), каноническая (рис. 1.13), последовательная (рис. 1.14), параллельная (рис. 1.15) и др.

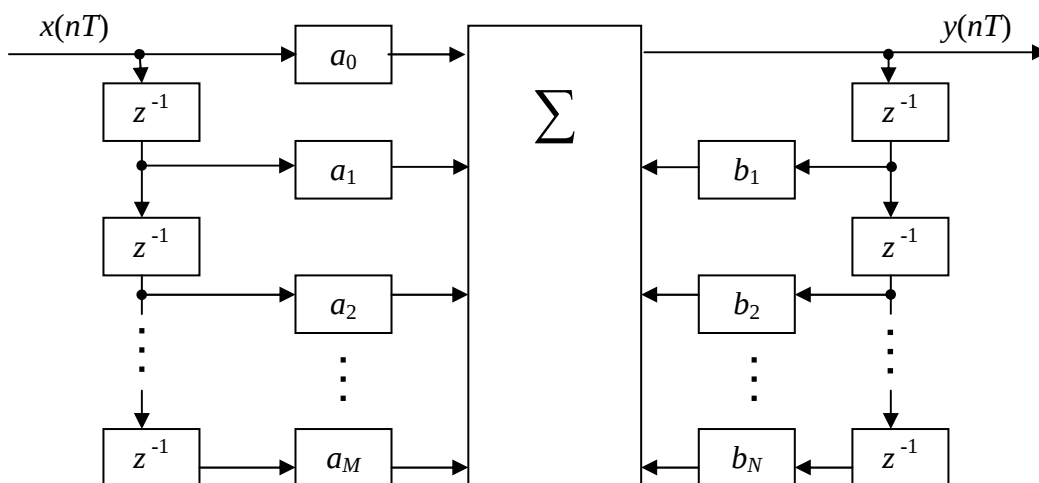


Рис. 1.12. – Основная (прямая) форма реализации ЛЦФ

Задержки (z^{-1}) необходимы для того, чтобы иметь доступ к предыдущим отсчетам сигнала.

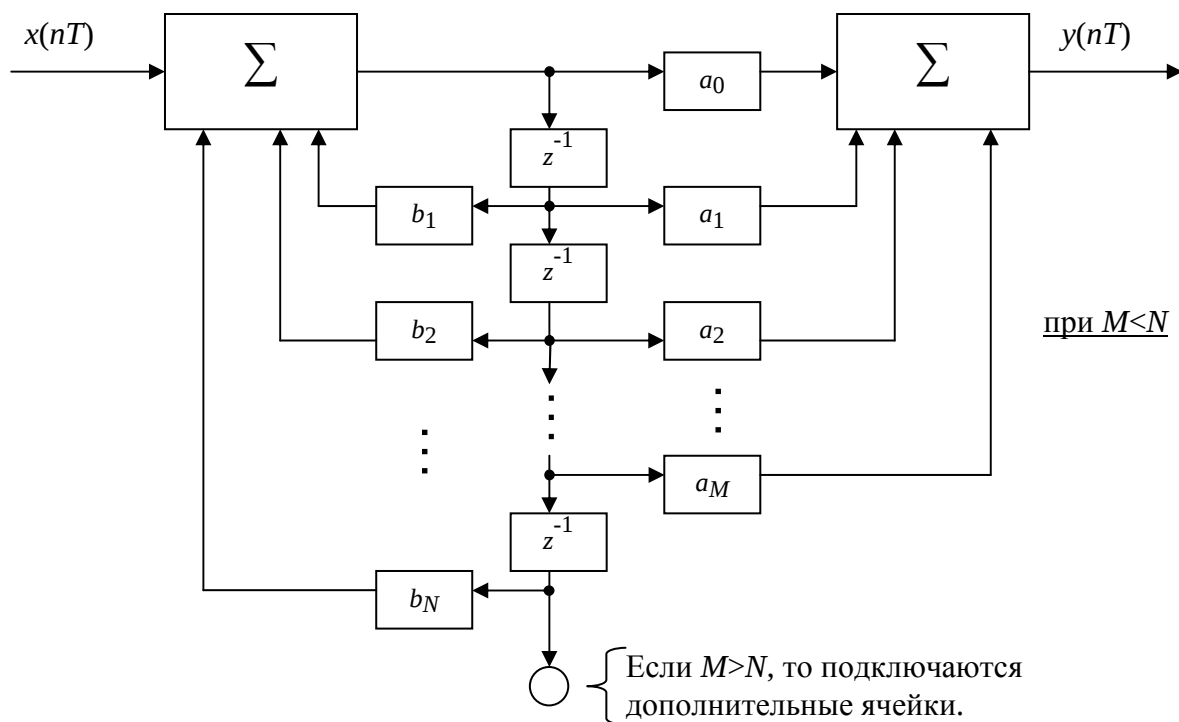


Рис. 1.13. – Каноническая форма реализации ЛЦФ (с минимальным количеством элементов задержки)

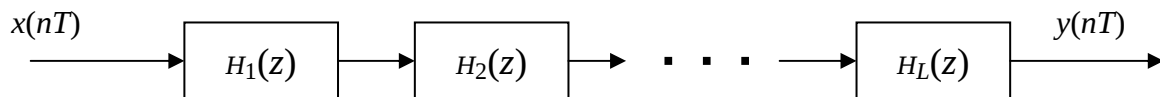


Рис. 1.14. – Последовательная (каскадная) форма реализации ЛЦФ

Здесь $H_{\Sigma}(z) = \prod_{l=1}^L H_l(z)$. Как правило, $H_l(z)$ есть ЦФ 1-го или 2-го порядков.

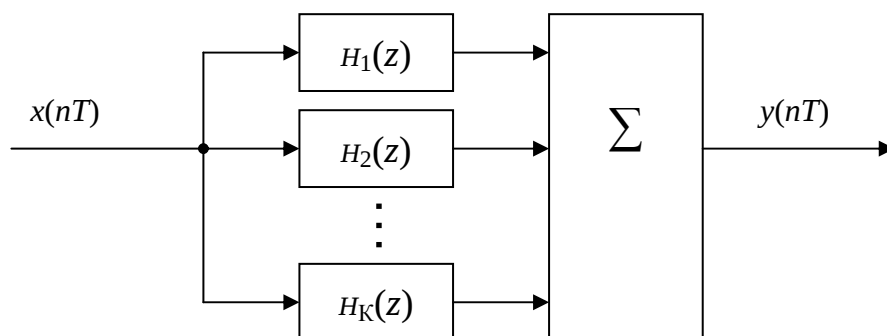


Рис. 1.15. – Параллельная форма реализации ЛЦФ

Для параллельной формы реализации ЛЦФ, приведенной на рис. 1.15,

$$H(z) = \sum_{k=1}^K H_k(z).$$

Существуют и другие формы реализации ЛЦФ (например, лестничная). Для дискретных фильтров, т.е. когда отсутствует операция квантования сигналов по уровню, выходной сигнал не зависит от формы реализации ЛЦФ.

Для цифровых фильтров, когда входные и промежуточные сигналы представлены квантованными по уровню дискретными последовательностями, различные формы реализации приводят к разным выходным сигналам.

Таким образом, погрешности фильтрации цифровых данных зависят как от разрядности представления входного и выходного массивов чисел, так и от последовательности выполнения операций над ними.

1.5. Характеристики ЛЦФ с постоянными параметрами

К характеристикам ЛЦФ с постоянными параметрами относятся:

- системная (передаточная) функция в Z-форме (СФ);
- импульсная и переходная характеристики (ИХ и ПХ, соответственно);
- комплексно-частотная, амплитудно-частотная и фазочастотная характеристики (КЧХ, АЧХ и ФЧХ, соответственно);
- групповое время запаздывания (ГВЗ);
- точностные характеристики (ТХ).

1.5.1. Системная (передаточная) функция фильтра в Z-форме

Системной (передаточной) функцией $H(z)$ фильтра называют отношение Z-образа выходного сигнала $Y(z)$ к Z-образу входного сигнала фильтра $X(z)$ при нулевых начальных условиях, т.е.

$$H(z) = \frac{Y(z)}{X(z)} \text{ при } y(-T) = y(-2T) = \dots = y(-NT) \equiv 0;$$

$$x(nT) \equiv 0 \text{ при } n < 0.$$

Применяя прямое Z-преобразование к левой и правой частям разностного уравнения, получим:

$$\begin{aligned} x(nT) &\rightarrow X(z); \\ x(nT-T) &\rightarrow X(z) \cdot z^{-1}; \\ x(nT-iT) &\rightarrow X(z) \cdot z^{-i}; \\ y(nT) &\rightarrow Y(z); \\ y(nT-T) &\rightarrow Y(z) \cdot z^{-1}; \end{aligned}$$

$$Y(z) = \sum_{i=0}^M a_i \cdot X(z) \cdot z^{-i} + \sum_{j=1}^N b_j \cdot Y(z) \cdot z^{-j};$$

$$Y(z) \cdot \left[1 - \sum_{j=1}^N b_j \cdot z^{-j} \right] = X(z) \cdot \sum_{i=0}^M a_i \cdot z^{-i};$$

$$H(z) = \frac{\sum_{i=0}^M a_i \cdot z^{-i}}{1 - \sum_{j=1}^N b_j \cdot z^{-j}}, \text{ где } z^{-1} = e^{-pT}.$$

В разностном уравнении и системной функции используются одни и те же коэффициенты, поэтому весьма просто выполнить переход от разностного уравнения к передаточной функции и структуре фильтра.

Разностное уравнение позволяет рассчитывать выходной сигнал фильтра лишь рекуррентно, т.е. для вычисления $y(nT)$ предварительно необходимо рассчитать $y(nT-T), \dots, y(nT-NT)$.

Зная системную функцию фильтра и входной сигнал (его Z-образ) при нулевых начальных условиях, можно получить в явном виде выражение для $y(nT)$ и его величину для любого наперед заданного значения $n = n_1$ без вычисления предыдущих отсчетов сигнала. Для этого:

а) $x(nT) \rightarrow X(z)$;

б) $Y(z) = X(z) \cdot H(z)$;

в) $y(nT) = \frac{1}{2\pi j} \oint_C Y(z) \cdot z^{n-1} dz \rightarrow y(nT) = Z^{-1}\{Y(z)\}.$

1.5.2. Импульсная и переходная характеристики

Импульсной характеристикой $h(nT)$ называют реакцию фильтра при нулевых начальных условиях на входное воздействие в виде единичного импульса $\delta(nT)$. При $x(nT) = \delta(nT) \rightarrow y(nT) = h(nT)$. В отличие от импульсной характеристики аналоговой цепи функция $h(nT)$ является безразмерной функцией.

Связь импульсной характеристики и системной функции

$$H(z) = \left. \frac{Y(z)}{X(z)} \right|_{x(nT)=\delta(nT)} = Y(z), \text{ так как } Z\{\delta(nT)\} = 1.$$

$H(z) = Y(z) = Z\{h(nT)\}$, т.е. Z-образ импульсной характеристики совпадает с системной функцией системы.

Дискретная свертка входного сигнала и импульсной характеристики фильтра:

$$y(nT) = x(nT) * h(nT) = \sum_{k=0}^n x(kT) \cdot h(nT - kT) = \sum_{k=0}^n x(nT - kT) \cdot h(kT),$$

где $n \geq 0$.

Дискретная свертка позволяет в явном виде получить выходной сигнал, зная входной сигнал $x(nT)$ и импульсную характеристику $h(nT)$.

На рис. 1.16 приведена иллюстрация формирования выходного сигнала в процессе свертки входного сигнала и импульсной характеристики фильтра.

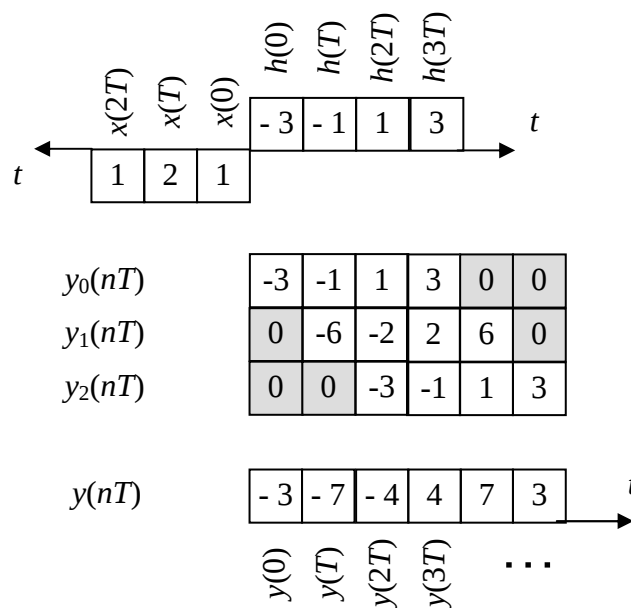


Рис 1.16 – Формирование выходного сигнала путем свертки входного сигнала с импульсной характеристикой фильтра (графическая интерпретация алгоритма)

Здесь сигнал $x(nT)$ «стоит на месте», а импульс $h(nT)$ «пробегают». При этом выходной сигнал можно найти как:

$$y(nT) = x(0)h(nT) + x(T)h(nT - T) + x(2T)h(nT - 2T) + \dots, n \geq 0.$$

Для приведенного случая $h(nT)$ представлена нечетной функцией из четырех отсчетов, а $x(nT)$ – в виде четной функции из трех отсчетов. Количество информационных отсчетов в выходном сигнале фильтра $y(nT)$ равно шести, т.е. сумме информативных отсчетов $x(nT)$ и $h(nT)$ минус один отсчет. На рисунке в каждый момент времени указано положение сигнала относительно импульсной характеристики, а отсчеты выходного сигнала формируются в соответствии с выражением для свертки после умножения соответствующих отсчетов $x(nT)$ и $h(nT)$ и их последующего суммирования. Выходной сигнал фильтра можно получить и сложением умноженных на значения каждого из

отсчетов входного сигнала отсчетов импульсной характеристики с соответствующими сдвигами во времени.

Важнейшей временной характеристикой цифрового фильтра является *переходная характеристика* (ПХ), под которой понимают реакцию фильтра $g(nT)$ на единичную ступенчатую функцию $u(nT)$ при нулевых начальных условиях:

$$g(nT) = y(nT) \Big|_{x(nT)=u(nT)}, \quad y(-T) = y(-2T) = \dots = y(NT) \equiv 0,$$

$$x(nT) \equiv 0 \text{ при } n < 0.$$

Связь изображения переходной характеристики $G(z)$ и системной функции $H(z)$:

$$X(z) = U(z) = Z\{u(nT)\} = \frac{z}{z-1},$$

$$G(z) = Y(z) = X(z)H(z) = H(z) \frac{z}{z-1}.$$

Таким образом:

$$G(z) = Z\{g(nT)\} = H(z) \frac{z}{z-1}, \quad g(nT) = Z^{-1}\left\{H(z) \frac{z}{z-1}\right\}.$$

Связь импульсной и переходной характеристик во временной области:

$$g(nT) = \sum_{k=0}^{\infty} h(nT - kT), \quad h(nT) = g(nT) - g(nT - T), \quad n \geq 0.$$

Анализ цифрового фильтра по переходной характеристике подобен анализу аналогового фильтра, т.е. по ПХ определяют время задержки фронта, время установления фронта, амплитуду первого выброса (перерегулирование), колебательность ПХ, спад вершины импульса и т.д.

1.5.3. Частотные характеристики цифровых фильтров

Пусть $X(e^{j\omega T})$ и $Y(e^{j\omega T})$ есть спектры дискретных последовательностей $x(nT)$ и $y(nT)$, т.е.

$$X(e^{j\omega T}) = \Phi\{x(nT)\} = \sum_{n=0}^{\infty} x(nT)e^{-j\omega nT}, \quad Y(e^{j\omega T}) = \Phi\{y(nT)\} = \sum_{n=0}^{\infty} y(nT)e^{-j\omega nT}.$$

Тогда комплексной частотной характеристикой (КЧХ) называют отношение $H(e^{j\omega T}) = \frac{Y(e^{j\omega T})}{X(e^{j\omega T})}$. То есть КЧХ совпадает с системной функцией

$H(z)$ на единичной окружности в z -плоскости, то есть при $z = e^{j\omega T}$: $H(e^{j\omega T}) = H(z) \Big|_{z=e^{j\omega T}}$.

На рис. 1.17, в z -плоскости, проиллюстрирована периодичность частотных характеристик ЦФ.

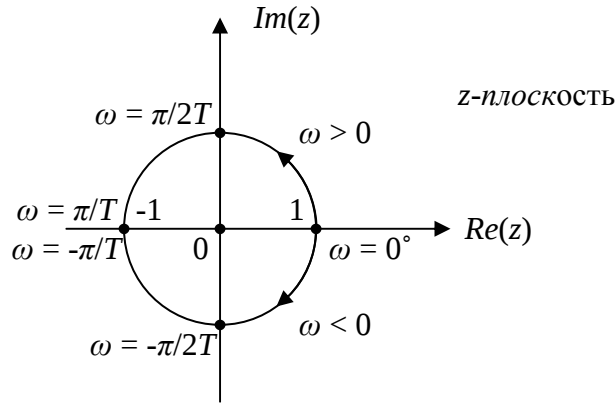


Рис. 1.17. – Периодичность КЧХ, АЧХ, ФЧХ ЦФ

Модуль КЧХ называют амплитудно-частотной характеристикой (АЧХ), т.е. $A(\omega) = |H(e^{j\omega T})|$ – АЧХ, а аргумент КЧХ есть фазочастотная характеристика (ФЧХ), т.е. $\varphi(\omega) = \arg[H(e^{j\omega T})]$.

Физический смысл КЧХ:

Для входного комплексного гармонического сигнала $x(nT) = e^{j\omega T}$ выходной сигнал в установившемся режиме ($n \rightarrow \infty$) имеет вид:

$$y(nT) = H(e^{j\omega T}) \cdot e^{j\omega nT}.$$

Иногда $H(e^{j\omega T})$ называют собственным значением (собственным числом), а $e^{j\omega nT}$ – собственной функцией.

Таким образом, если на вход фильтра подан дискретный сигнал в виде дискретизированной синусоиды, то $A(\omega)$ характеризует амплитуду выходного сигнала, а $\varphi(\omega)$ – его фазу, т.е. подавая на вход линейного цифрового фильтра гармонический сигнал, на его выходе в установившемся режиме ($n \rightarrow \infty$) получим гармонический сигнал той же частоты, но с другой амплитудой и фазой, т.к. в линейных системах *новых частот не возникает*.

Для вещественных a_i, b_j АЧХ и ФЧХ можно вычислить по формулам:

$$\begin{aligned} A(\omega) = |H(e^{j\omega T})| &= \frac{\sqrt{\left[\sum_{i=0}^M a_i \cos(i\omega T) \right]^2 + \left[\sum_{i=0}^M a_i \sin(i\omega T) \right]^2}}{\sqrt{\left[\sum_{j=0}^N b_j \cos(j\omega T) \right]^2 + \left[\sum_{j=0}^N b_j \sin(j\omega T) \right]^2}} = \\ &= \sqrt{\frac{\sum_{m=0}^M \sum_{k=0}^M a_m a_k \cos[(m-k)\omega T]}{\sum_{p=0}^N \sum_{s=0}^N b_p b_s \cos[(p-s)\omega T]}}; \end{aligned}$$

$$\varphi(\omega) = \operatorname{arctg} \frac{\operatorname{Im}[H(e^{j\omega T})]}{\operatorname{Re}[H(e^{j\omega T})]} = -\operatorname{arctg} \frac{\sum_{i=0}^M a_i \sin(i\omega T)}{\sum_{i=0}^M a_i \cos(i\omega T)} + \operatorname{arctg} \frac{\sum_{j=0}^N b_j \sin(j\omega T)}{\sum_{j=0}^N b_j \cos(j\omega T)}, \quad b_0 = -1.$$

Пример:

Пусть $H(z) = \frac{1}{1 - b_1 z^{-1}}$; $b_1 = e^{-\frac{T}{\tau}}$, тогда

$$\text{АЧХ: } A(\omega) = \frac{1}{\sqrt{[1 - b_1 \cos(\omega T)]^2 + b_1^2 \sin^2(\omega T)}} = \frac{1}{\sqrt{1 + b_1^2 - 2b_1 \cos(\omega T)}};$$

$$\text{ФЧХ: } \varphi(\omega) = \operatorname{arctg} \frac{b_1 \sin(\omega T)}{b_1 \cos(\omega T) - 1} = \operatorname{arctg} \frac{\sin(\omega T)}{\cos(\omega T) - b_1^{-1}}.$$

1.5.4. Групповое время запаздывания

Групповое время запаздывания (ГВЗ) является мерой средней задержки узкополосных сигналов в инерционных линейных цепях, в том числе и в цифровых фильтрах и определяется следующим образом:

$$\tau_{gp}(\omega) = -\frac{d\varphi(\omega)}{d\omega}.$$

Предпочтительна приблизительно постоянная характеристика групповой задержки во всей полосе пропускания цифрового фильтра.

Для выравнивания времени запаздывания по полосе частот применяют фазовые корректоры на основе неминимально-фазовых звеньев, имеющих постоянную АЧХ в полосе пропускания ($A(\omega) = \text{const}$) и изменяющуюся фазовую характеристику ($\varphi(\omega) = \text{var}$). Характерным признаком неминимально-фазовых звеньев является наличие нулей вне единичной окружности в z -плоскости [12].

1.5.5. Устойчивость цифровых фильтров

Фильтр называется устойчивым, если при любых начальных условиях реакция фильтра 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 , причем M_x и M_y – постоянные, не зависящие от n . Из уравнения дискретного фильтра следует, что если $x(nT)$ ограничено, т.е. $|x(nT)| \leq M_x < \infty$ для всех n , то абсолютное значение выходного сигнала

$$|y(nT)| \leq \sum_{m=0}^{\infty} |h(mT)| \cdot |x(nT - mT)| \leq M_x \sum_{m=0}^{\infty} |h(mT)|.$$

Следовательно, критерием устойчивости дискретного фильтра является абсолютная сходимость ряда отсчетов импульсной характеристики

$$\sum_{n=0}^{\infty} |h(nT)| < \infty.$$

Можно показать, что данное условие является не только достаточным, но и необходимым условием устойчивости фильтра. Однако непосредственное применение этого условия для проверки устойчивости практически затруднено. Поэтому рассмотрим другую формулировку критерия устойчивости.

В общем случае передаточная функция линейного дискретного фильтра представляется в виде $H(z) = \sum_{n=0}^{\infty} h(nT)z^{-n}$, поэтому $|H(z)| \leq \sum_{n=0}^{\infty} |h(nT)| \cdot |z^{-n}|$.

Если $|z^{-1}| \leq 1$, то $H(z) = \sum_{n=0}^{\infty} |h(nT)|$. Это значит, что в *устойчивой системе* $H(z)$ конечна во всех точках z -плоскости, где $|z| \geq 1$, и, следовательно, передаточная функция $H(z)$ не должна иметь особых точек (полюсов) при $z \geq 1$ (на и вне единичного круга z -плоскости). Таким образом, система будет устойчива только тогда, когда все полюсы $H(z)$ расположены внутри единичного круга z -плоскости. Если $H(z)$ – дробно-рациональная функция, то полюсы $H(z)$ определяются нулями (корнями) многочлена знаменателя передаточной функции:

$$H(z) = \frac{\sum_{i=0}^M a_i z^{-i}}{1 - \sum_{j=1}^N b_j z^{-j}}.$$

Пример:

1. Пусть $H(z) = \frac{1 - z^{-1}}{1 - 0,2z^{-1}}$. Полюс $H(z)$ – корень знаменателя $z_1 = 0,2 < 1$;

фильтр устойчив.

2. Пусть $H(z) = \frac{1 - z^{-3}}{1 - 0,6z^{-1} + 0,25z^{-2}}$; здесь корни знаменателя

$z_{2,3} = 0,3 \pm j0,4$ и $|z_2| = |z_3| = 0,5 < 1$, т.е. фильтр устойчив.

3. Пусть $H(z) = \frac{1}{1 + 1,2z^{-1}}$; здесь корень знаменателя $z_4 = -1,2$, т.е. $|z_4| > 1$,

следовательно, фильтр неустойчив.

Заметим, что формулировка критерия устойчивости относится к несократимой дроби, так как возможно, что полюс компенсируется нулем передаточной функции. Например, фильтр, описываемый функцией $H(z) = \frac{1 - z^{-2}}{1 - z^{-1}}$,

устойчив, несмотря на то, что полюс $|z| = 1$ лежит на единичной окружности; устойчивость фильтра обусловлена тем, что в действительности после сокращения числителя и знаменателя на множитель $1 - z^{-1}$ получается $H(z) = 1 + z^{-1}$. Нерекурсивные фильтры всегда устойчивы, так как все полюсы их передаточной функции локализованы в бесконечно удаленной точке ($|z| = \infty$).

Алгебраический критерий устойчивости рекурсивного ЦФ

Задачу об устойчивости рекурсивного ЦФ произвольного порядка можно решить, связав данную проблему с расположением полюсов в w -плоскости.

Для этого заметим, что преобразование вида $z = \frac{w+1}{w-1}$ взаимно-однозначно отображает левую полуплоскость комплексной переменной w на единичный круг в комплексной плоскости Z с центром в точке $z = 0$. Действительно, точке $w = -1$ соответствует точка $z = 0$. В то же время мнимая ось в w -плоскости, т.е. совокупность точек с координатами $w = ja$ (a – произвольное вещественное число), отображается в множество точек единичной окружности $z = -e^{j2\arctg(a)}$.

Возьмем характеристическое уравнение ЦФ $z^n - b_1 z^{n-1} - \dots - b_n = 0$ и подставим в него переменную z , выраженную через переменную w :

$$\left(\frac{w+1}{w-1}\right)^n - b_1 \left(\frac{w+1}{w-1}\right)^{n-1} - b_2 \left(\frac{w+1}{w-1}\right)^{n-2} - \dots - b_n = 0.$$

Приведя это выражение к общему знаменателю $(w-1)^n$, получим характеристическое уравнение относительно переменной w :

$$(w+1)^n - b_1(w-1)(w+1)^{n-1} - b_2(w-1)^2(w+1)^{n-2} - \dots - b_n(w-1)^n = 0.$$

Если многочлен по степеням w , образующий левую часть последней формулы, имеет корни лишь в левой полуплоскости, то исходный характеристический многочлен имеет корни, располагающиеся лишь в единичном круге z -плоскости. Как следствие, анализируемый рекурсивный ЦФ будет устойчивым.

Отметим, что многочлен $(w-1)^n$ не имеет корней в левой плоскости. Поэтому приведение к общему знаменателю правомерно.

Критерий устойчивости Рауса-Гурвица

Данный критерий устойчивости относится к числу алгебраических критериев, т.е. критериев, сформулированных в виде некоторых алгебраических формул. По критерию Рауса-Гурвица анализируются коэффициенты характеристического уравнения:

$$a_0 w^n + a_1 w^{n-1} + a_2 w^{n-2} + \dots + a_n = 0.$$

1. Делаем коэффициент положительным, т.е. $a_0 > 0$. Если $a_0 < 1$, то умножаем левую часть характеристического уравнения на минус единицу (-1) .

2. Составим из коэффициентов характеристического уравнения квадратную матрицу, следуя определенным правилам.

По главной диагонали матрицы коэффициенты уравнения от a_1 до a_n . Затем каждый столбец матрицы заполняют коэффициентами этого же уравнения: вверх – в порядке возрастания индекса коэффициентов, вниз – в порядке убывания. В тех местах каждого столбца, где индекс оказывается отрицательным или превышает n , записывают нуль. Таким образом, матрица имеет вид:

$$\begin{bmatrix} a_1 & a_3 & a_5 & \dots & 0 & 0 \\ a_0 & a_2 & a_4 & \dots & 0 & 0 \\ 0 & a_1 & a_3 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_{n-1} & 0 \\ 0 & 0 & 0 & \dots & a_{n-2} & a_n \end{bmatrix}.$$

Затем из элементов этой матрицы, расположенных симметрично относительно главной диагонали, составляют определители:

$$\Delta_1 = a_1, \Delta_2 = \begin{vmatrix} a_1 & a_3 \\ a_0 & a_2 \end{vmatrix}, \Delta_3 = \begin{vmatrix} a_1 & a_3 & a_5 \\ a_0 & a_2 & a_4 \\ 0 & a_1 & a_3 \end{vmatrix}, \dots,$$

$$\Delta_{n-1} = \begin{vmatrix} a_1 & a_3 & a_5 & \dots & 0 \\ a_0 & a_2 & a_4 & \dots & 0 \\ 0 & a_1 & a_3 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & a_{n-1} \end{vmatrix}, \Delta_n = \begin{vmatrix} a_1 & a_3 & a_5 & \dots & 0 & 0 \\ a_0 & a_2 & a_4 & \dots & 0 & 0 \\ 0 & a_1 & a_3 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_{n-1} & 0 \\ 0 & 0 & 0 & \dots & a_{n-2} & a_n \end{vmatrix}.$$

Критерий формулируется следующим образом: для устойчивости системы необходимо и достаточно, чтобы при $a_0 > 0$ все n определителей были положительны.

Отметим, что $\Delta_1 = a_1 > 0$ в силу необходимого условия устойчивости, а $\Delta_n = a_n \Delta_{n-1}$, т.к. последний столбец определителя Δ_n содержит лишь один ненулевой элемент a_n , причем $a_n > 0$. Таким образом, объединяя условия 1 и 2, получим: $a_0 > 0, a_1 > 0, \dots, a_n > 0, \Delta_k > 0, k = 2, \dots, n-1$.

Условия устойчивости фильтра 2-го порядка.

$$a_n > 0, n = 0, 1, 2; \Delta_1 = a_1 > 0, \Delta_2 = \begin{vmatrix} a_1 & 0 \\ a_0 & a_2 \end{vmatrix} = a_1 a_2 - 0 \cdot a_0 = a_1 a_2.$$

Условия устойчивости фильтра 3-го порядка.

$$a_0 > 0, a_1 > 0, a_3 > 0, \Delta_1 = a_1 > 0, \Delta_2 = \begin{vmatrix} a_1 & a_3 \\ a_0 & a_2 \end{vmatrix} = a_1 a_2 - a_0 a_3 > 0,$$

$$\Delta_3 = \begin{vmatrix} a_1 & a_3 & 0 \\ a_0 & a_2 & 0 \\ 0 & a_1 & a_3 \end{vmatrix} = a_3 \Delta_2 = a_3 (a_1 a_2 - a_0 a_3) > 0.$$

Окончательно: $a_0 > 0, a_1 > 0, a_3 > 0, a_1 a_2 - a_0 a_3 > 0$.

Условия устойчивости фильтра 4-го порядка.

$$a_0 > 0, a_1 > 0, a_4 > 0, \Delta_1 = a_1 > 0, \Delta_2 = \begin{vmatrix} a_1 & a_3 \\ a_0 & a_2 \end{vmatrix} = a_1 a_2 - a_0 a_3 > 0,$$

$$\Delta_3 = \begin{vmatrix} a_1 & a_3 & 0 \\ a_0 & a_2 & a_4 \\ 0 & a_1 & a_3 \end{vmatrix} = a_3 (a_1 a_2 - a_0 a_3) - a_1^2 a_4 > 0,$$

$$\Delta_4 = \begin{vmatrix} a_1 & a_3 & 0 & 0 \\ a_0 & a_2 & a_4 & 0 \\ 0 & a_1 & a_3 & 0 \\ 0 & a_0 & a_2 & a_4 \end{vmatrix} = a_4 \Delta_3 > 0.$$

Окончательно: $a_0 > 0, a_1 > 0, a_4 > 0, a_1 a_2 - a_0 a_3 > 0,$
 $a_3 (a_1 a_2 - a_0 a_3) - a_1^2 a_4 > 0.$

Пример:

Исследовать устойчивость рекурсивного цифрового фильтра 3-го порядка с характеристическим уравнением $z^3 + 0,4z^2 - 0,5z + 1 = 0$.

В соответствие с формулой получаем преобразованное уравнение:

$$(w+1)^3 + 0,4(w-1)(w+1)^2 - 0,5(w-1)^2(w+1) + 1 =$$

$$= 0,9w^3 + 3,9w^2 + 3,1w + 1,1 = 0.$$

Здесь все коэффициенты положительны и в то же время

$a_1 a_2 - a_0 a_3 = 11,1 > 0$. По критерию Рауса-Гурвица данный многочлен устойчив. Значит устойчив и анализируемый цифровой фильтр.

1.5.6. Точностные характеристики цифровых фильтров

Все погрешности, связанные с дискретизацией по уровню и квантованием результатов арифметических операций, ухудшают точность фильтрации, поэтому встают задачи оценки и минимизации этих погрешностей.

Какие же погрешности возникают в цифровых фильтрах?

Это прежде всего:

1. Погрешности, возникающие при аппроксимации заданных частотных и временных характеристик проектируемого цифрового фильтра (определяется порядок фильтра и коэффициенты a_i, b_j).
2. Погрешности, связанные с квантованием входного сигнала по уровню в АЦП (шумы квантования).
3. Погрешности, связанные с квантованием коэффициентов цифрового фильтра a_i и b_j .
4. Погрешности, связанные с квантованием результатов арифметических операций (особенно при умножении и делении).
5. Эффекты, связанные с переполнением разрядной сетки арифметического устройства (сумматора) цифрового фильтра.

Существует два подхода к оценке погрешностей:

- детерминированный подход;
- вероятностный подход.

При детерминированном подходе оцениваются максимально возможные значения ошибок, которые могут возникнуть при вычислениях. Его используют для оценки 5-ой погрешности (см. выше список погрешностей), возникающей при переполнении разрядной сетки регистров или ячеек памяти. Детерминированный подход требует либо моделирования, либо непосредственного расчета разрядностей. Детерминированная оценка, как правило, дает пессимистический прогноз. Если вероятность ошибки мала, то мы можем сильно исказить результат.

Вероятностный подход используется при оценке эффектов квантования и округления результатов арифметических операций, т.е. для оценки 2-ой и 4-ой погрешностей (см. выше список погрешностей).

Обозначим источники погрешностей через $e_k(nT)$, где $k = 0, 1, 2, 3, \dots$. При $k = 0$ оценивается источник шума квантования по уровню в АЦП; при $k = 1, 2, 3, \dots$ описываются шумы, возникающие в арифметическом устройстве.

При вероятностном подходе используются следующие предположения:

1. Отсчеты шума от всех источников погрешностей $e_k(nT)$ представляются дискретными белыми шумами с равномерным законом распределения и дисперсией σ_k^2 .

2. Все источники шума не коррелированы между собой.
3. Любой из источников шума $e_k(nT)$ не коррелирован с входным сигналом.

Достоинство вероятностной оценки: по аналитическим формулам можно определить суммарную погрешность $\sigma_{\sum_{\text{вых}}}^2 = \sum_{k=0}^K \sigma_{k_{\text{вых}}}^2$, здесь $\sigma_{0_{\text{вых}}}^2$ – погрешность, связанная с шумом квантования АЦП, $\sigma_{k_{\text{вых}}}^2$ – дисперсии остальных погрешностей. Используя принцип суперпозиции можно оценивать и минимизировать каждую компоненту порознь.

Однако, чтобы не абсолютизировать вероятностную оценку, необходимо помнить, что используемые предположения в реальных цифровых фильтрах не всегда соблюдаются. В частности, появление предельных циклов обусловлено коррелированным шумом округления.

В цифровых фильтрах может использоваться как округление, так и усечение результатов (рис. 1.18).

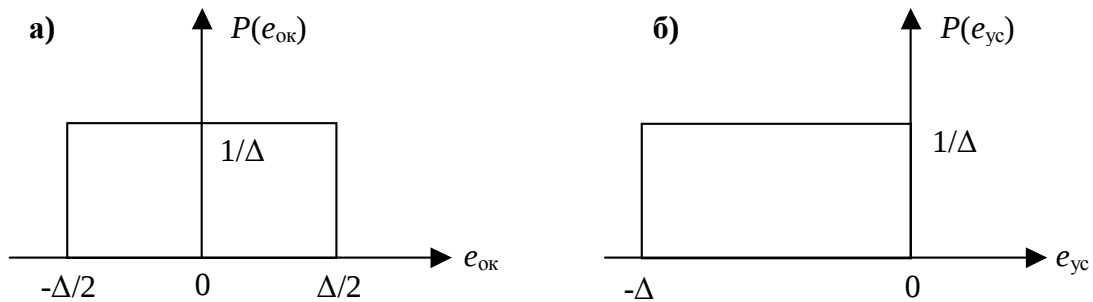


Рис. 1.18. – Плотность вероятности ошибок квантования:
а) округление данных;
б) усечение данных.

Например:

$$\left. \begin{array}{l} 6,4 + 0,5 \rightarrow 6 \\ 6,6 + 0,5 \rightarrow 7 \end{array} \right\} - \text{округление, а} \quad \left. \begin{array}{l} 6,4 \rightarrow 6 \\ 6,6 \rightarrow 6 \end{array} \right\} - \text{усечение.}$$

При этом, используя формулы для математического ожидания $\mu_e = E\{e\} = \int_{-\infty}^{\infty} e \cdot p(e) de$ и дисперсии $\sigma_e^2 = E\{e^2\} = \int_{-\infty}^{\infty} e^2 \cdot p(e) de$ (при $\mu_e = 0$) и

$\sigma_e^2 = E\{e^2\} = \mu_e^2 = \int_{-\infty}^{\infty} (e - \mu_e)^2 \cdot p(e) de$ (при $\mu_e \neq 0$), получим:

$$\mu_{e_{\text{окр}}} = 0, \quad \sigma_{e_{\text{окр}}}^2 = \frac{\Delta^2}{12} - \text{при округлении;}$$

$$\mu_{e_{\text{ус}}} = -\frac{\Delta}{2}, \quad \sigma_{e_{\text{ус}}}^2 = \frac{\Delta^2}{12} - \text{при усечении.}$$

Математические ожидания при округлении и усечении данных различны, а дисперсии равны.

При анализе точностных характеристик ЦФ можно говорить о локальных системных функциях $H_k(z)$ и локальных импульсных характеристиках $h_k(nT)$, для каждого из источников погрешностей $e_k(nT)$ (рис. 1.19).

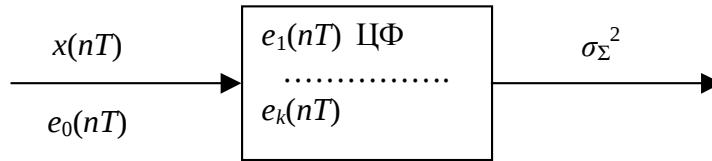


Рис. 1.19. – Погрешности, влияющие на точность цифровой фильтрации

Если мы определим точки возникновения погрешности и их функции, то каждую из компонент погрешности можно вычислить по формулам:

1. $\sigma_{квх}^2 = \sigma_{квх}^2 \sum_{n=0}^{\infty} h_k^2(nT)$, $\sigma_{0вх}^2 = \frac{\Delta^2}{12} \sum_{n=0}^{\infty} h^2(nT)$ – по локальной ИХ;
2. $\sigma_{квх}^2 = \sigma_{квх}^2 \frac{1}{2\pi j} \oint H_k(z) \cdot H_k(z^{-1}) \cdot z^{-1} dz$ – по локальной системной функции;
3. $\sigma_{квх}^2 = \sigma_{квх}^2 \frac{T}{\pi} \int_0^{\pi/T} |H_k(e^{j\omega T})|^2 d\omega$ – по локальным АЧХ.

По любой из формул можно вычислить $\sigma_{квх}^2$ и затем просуммировать. Для расчета погрешностей на ЭВМ, а также для анализа точности нерекурсивных цифровых фильтров лучше всего подходит первая формула. Однако, если есть возможность использовать символьный процессор, то лучше использовать вторую формулу.

В случае возникновения каких-либо трудностей с выполнением аналитических расчетов можно произвести оценку моделированием, как показано на рисунке 1.20.

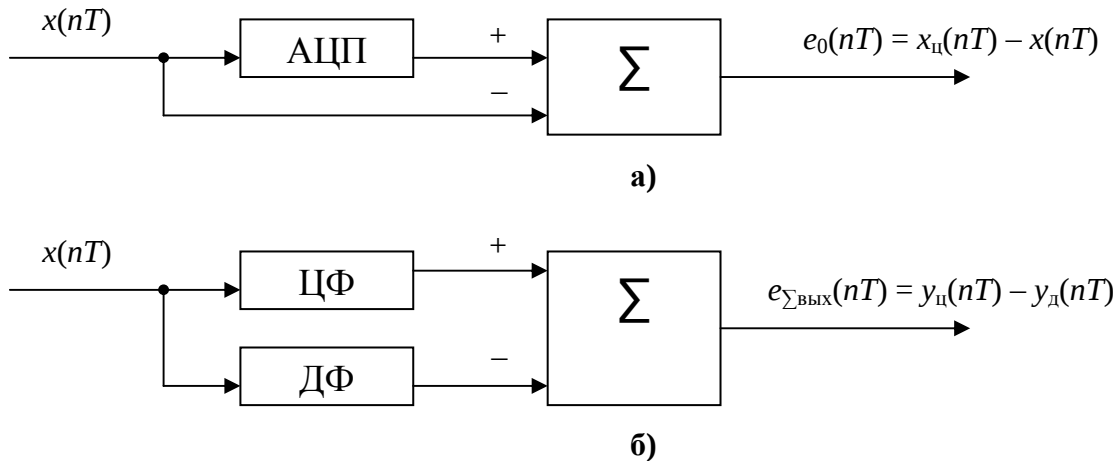


Рис. 1.20. – а) – выделение ошибки квантования в АЦП, б) – выделение суммарной ошибки ЦФ

Здесь ДФ – дискретный фильтр, т.е. фильтр без погрешностей (в нем происходят точные вычисления).

При использовании метода, показанного на рисунке 1.20б, исключается погрешность АЦП $e_0(nT)$.

Следует учесть, что при вероятностном подходе пятая погрешность становится слишком оптимистичной и в этом случае возможно переполнение разрядной сетки ЦФ.

1.6. Примеры построения линейных ЦФ

Цифровой дифференциатор или вычислитель первой разности (В1Р)

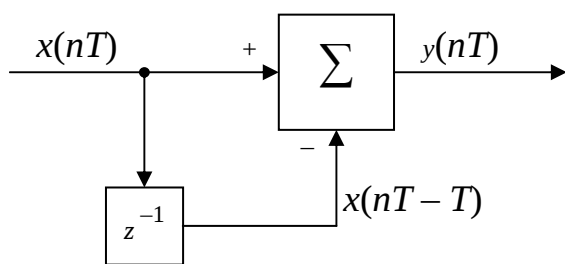


Рис. 1.21. – Структурная схема цифрового дифференциатора

Фильтр линейный и нерекурсивный (трансверсальный) (рис. 1.21).

РУ: $y(nT) = x(nT) - x(nT - T), n \geq 0$.

Здесь $a_0 = 1; a_1 = -1; b_j = 0$.

Системная функция:

$$Y(z) = X(z) - X(z)z^{-1} = X(z)(1 - z^{-1}),$$

$$H(z) = \frac{Y(z)}{X(z)} = 1 - z^{-1}.$$

КЧХ:

$$H(e^{j\omega T}) = 1 - e^{j\omega T} = 1 - \cos(\omega T) + j \cdot \sin(\omega T);$$

АЧХ (рис. 1.22):

$$A(\omega) = |H(e^{j\omega T})| = \sqrt{(1 - \cos(\omega T))^2 + \sin^2(\omega T)} = \sqrt{2 - 2\cos(\omega T)} = 2 \left| \sin\left(\frac{\omega T}{2}\right) \right|;$$

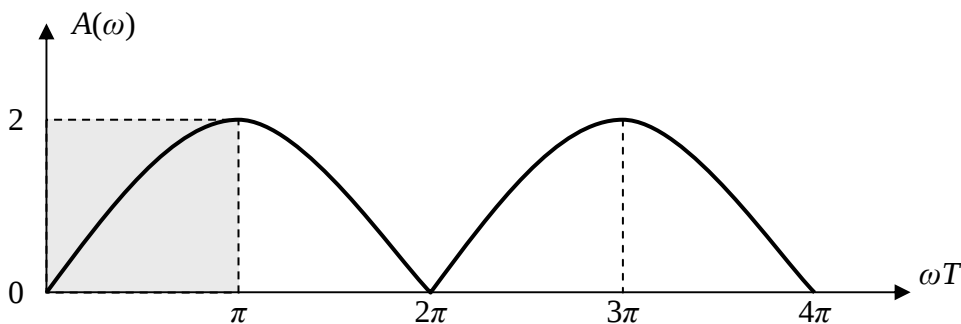


Рис. 1.22. – Амплитудно-частотная характеристика

$\Phi\mathcal{X}$ (рис. 1.23):

$$\begin{aligned}\varphi(\omega) &= \arg[H(e^{j\omega T})] = \arctg\left(\frac{\sin(\omega T)}{1 - \cos(\omega T)}\right) = \\ &= \arctg\left\{\frac{2\sin\left(\frac{\omega T}{2}\right)\cos\left(\frac{\omega T}{2}\right)}{1 - \left(1 - 2\sin^2\left(\frac{\omega T}{2}\right)\right)}\right\} = \frac{\pi - \omega T}{2}, \quad \omega \in \left[0, \frac{2\pi}{T}\right].\end{aligned}$$

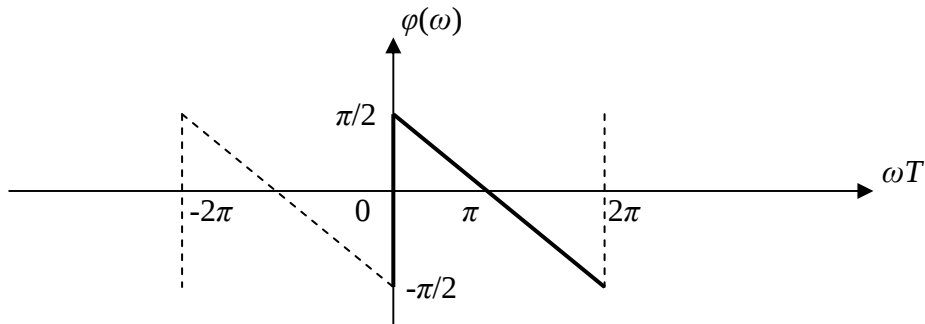


Рис. 1.23. – Фазочастотная характеристика

ГВЗ: $\tau(\omega) = -\frac{d\varphi(\omega)}{d\omega} = \frac{T}{2}.$

Если на вход дифференциатора подать единичный импульс $\delta(nT)$, то импульсная характеристика может быть записана в аналитическом виде как

ИХ (рис. 1.24):

$$h(nT) = y(nT)\big|_{x(nT)=\delta(nT)} = \delta(nT) - \delta(nT - T), \quad n \geq 0.$$

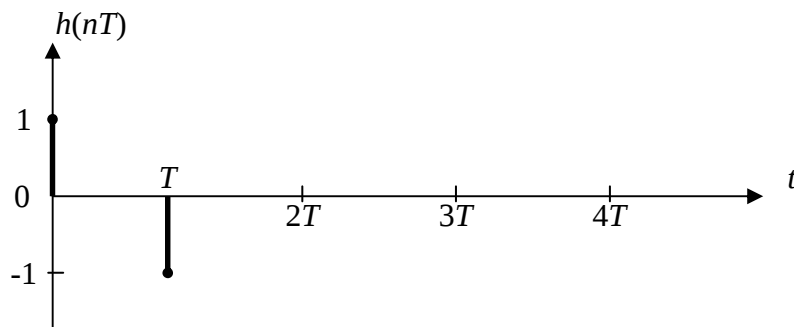


Рис. 1.24. – Импульсная характеристика (ИХ)

Пример прохождения сигнала через цифровой дифференциатор

Пусть $x(nT) = u(nT) - u(nT - 5T)$, $n \geq 0$, тогда $y(nT) = \delta(nT) - \delta(nT - 5T)$ (рис. 1.25):

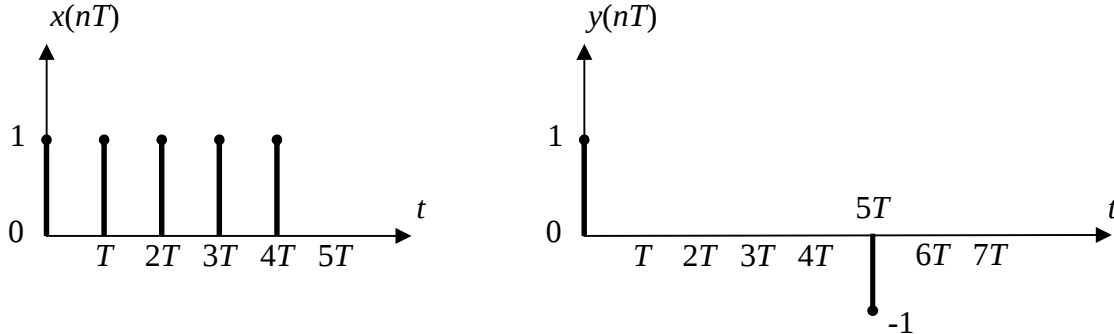


Рис. 1.25. – Входной и выходной сигналы

Шумы квантования на выходе цифрового дифференциатора (ЦД)

Используем рассмотренный выше пример (рис. 1.21). Пусть Δ – шаг квантования, а внесение шумов в сигнал представим, как на рис. 1.26:

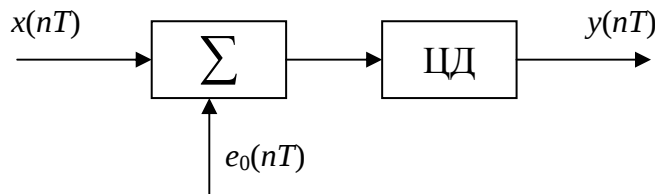


Рис. 1.26. – Внесение шума квантования АЦП

В этом случае $\sigma_{\text{вых}}^2 = \sigma_{\text{вх}}^2 \sum_{n=0}^{\infty} h^2(nT) = 2\sigma_{\text{вх}}^2 = \frac{\Delta^2}{6}$, т.е. дисперсия шума на выходе ЦД увеличивается в 2 раза по сравнению с входной.

Вычислитель второй разности (В2Р) (рис. 1.27)

Структурная схема вычислителя второй разности может быть получена последовательным включением двух вычислителей первой разности:

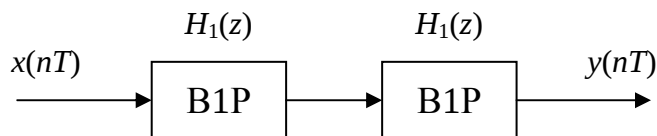


Рис. 1.27. – Вычислитель второй разности

СФ: $H(z) = H_1(z)H_1(z) = (1 - z^{-1})(1 - z^{-1}) = 1 - 2z^{-1} + z^{-2}$.

Тогда Z-изображение выходного сигнала будет выглядеть как:

$Y(z) = X(z) - 2X(z)z^{-1} + X(z)z^{-2}$. Применим формально обратное Z-преобразование и получим выражение для выходного сигнала:

$$y(nT) = x(nT) - 2x(nT - T) + x(nT - 2T), \quad n \geq 0.$$

На рис. 1.28 приведена структурная схема В2Р и его ИХ.

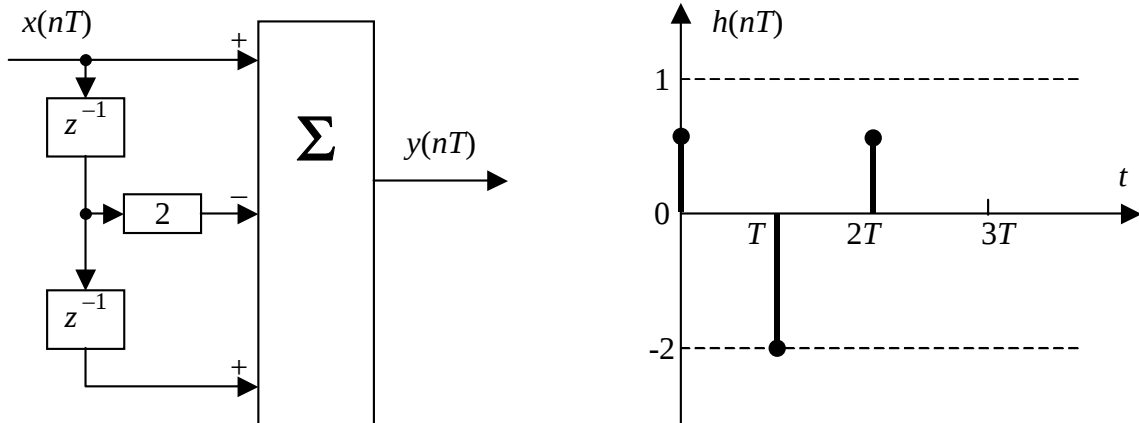
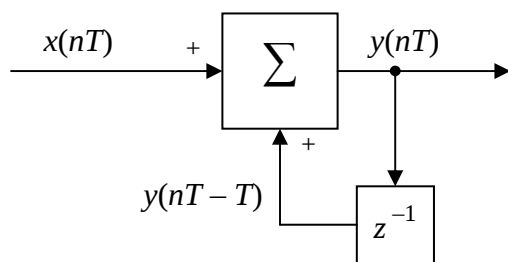


Рис. 1.28. – Вычислитель второй разности и его импульсная характеристика

Дисперсия шума на выходе вычислителя второй разности равна:

$$\sigma_{\text{вых}}^2 = (1 + 4 + 1)\sigma_{\text{вх}}^2 = \frac{\Delta^2}{2}.$$

Цифровой интегратор (накапливающий сумматор) (рис. 1.29)



$$PY: \quad y(nT) = x(nT) + y(nT - T)$$

Системная функция:

$$Y(z)(1 - z^{-1}) = X(z) \Rightarrow H(z) = \frac{1}{1 - z^{-1}} = \frac{z}{z - 1}.$$

Используя обратное Z-преобразование, найдем импульсную характеристику:

$$h(nT) = Z^{-1}\{H(z)\} = u(nT), \quad \text{т.е.} \quad \text{единичный скачок (см. таблицу 1.3).}$$

Рис. 1.29. – Структурная схема цифрового интегратора

Пусть на входе интегратора сигнал имеет вид $x(nT) = u(nT) - u(nT - 5T)$, тогда на выходе сигнал будет иметь вид, как на рис. 1.30:

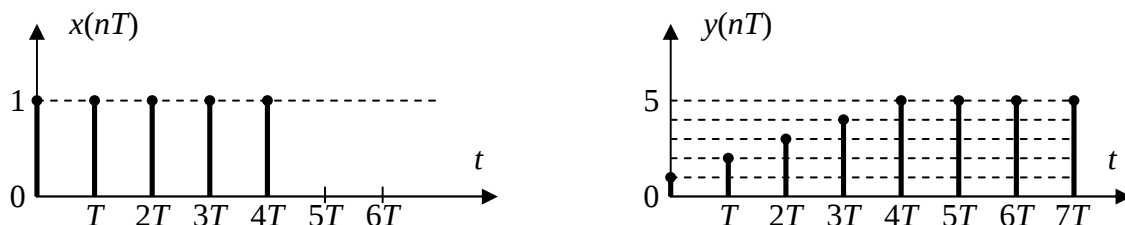


Рис. 1.30. – Сигналы на входе и выходе цифрового интегратора

Цифровой интегратор находится на границе устойчивости. Критерием устойчивости ЦФ является абсолютная сходимость ряда отсчетов импульсной характеристики:

$$\sum_{n=0}^{\infty} |h(nT)| < \infty.$$

Для обеспечения устойчивости интегратора необходимо ограничивать время суммирования (количество отсчетов).

Дисперсия выходного шума цифрового интегратора:

$$\sigma_{\text{вых}}^2 = \sigma_{\text{вх}}^2 \sum_{n=0}^M h^2(nT) = M\sigma_{\text{вх}}^2, \text{ т.е. при увеличении числа отсчетов } (M \rightarrow \infty)$$

дисперсия выходного шума увеличивается бесконечно, но при конечном M она также конечна.

Цифровой интегратор с ограниченным временем суммирования (рис. 1.31)

Ограниченное число отсчетов ИХ в ЦФ можно реализовать следующим образом (рекурсивный цифровой фильтр):

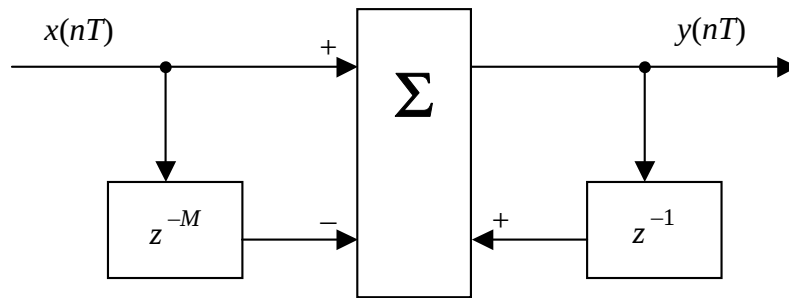


Рис. 1.31. – Цифровой интегратор с ограниченным временем суммирования

ИХ такого фильтра будет иметь вид: $h(nT) = u(nT) - u(nT - MT)$, $n \geq 0$. При этом выходной сигнал $y(nT) = x(nT) - x(nT - MT) + y(nT - 1)$, т.е. если подать на вход такого фильтра сигнал вида $x(nT) = \begin{cases} \delta(nT), & n = 0, \\ 0, & n \neq 0, \end{cases}$

то на выходе фильтра сигнал будет иметь вид, представленный на рис. 1.32.

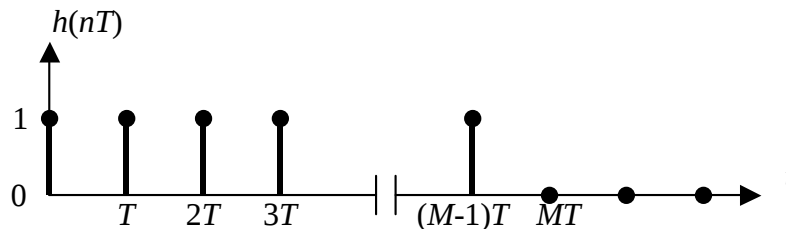


Рис. 1.32. – Импульсная характеристика ЦИ с ограниченным временем суммирования

Применяя к ИХ прямое Z-преобразование, получим выражение для системной функции $H(z)$:

$$H(z) = Z\{h(nT)\} = \sum_{n=0}^{\infty} h(nT)z^{-n} = \sum_{n=0}^{M-1} z^{-n} = \frac{1-z^{-M}}{1-z^{-1}} \quad \text{— это следует из формулы для суммы ряда ограниченной геометрической прогрессии, т.е.}$$

мулы для суммы ряда ограниченной геометрической прогрессии, т.е.

$$\sum_{n=0}^{M-1} q^{-n} = \frac{1-q^M}{1-q}, \quad \text{при } |q| < 1.$$

Сглаживающий фильтр (СФ) на базе цифрового интегратора (рис. 1.33)

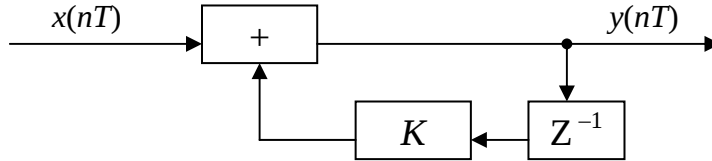


Рис. 1.33. – Цифровой сглаживающий фильтр

При $-1 < K < 0$ – ВЧ фильтр; при $K = 1$ – цифровой интегратор; при $|K| < 1$ – сглаживающий фильтр, а при $|K| \geq 1$ – неустойчивый фильтр, т.е. при любом малом возмущении $x(nT)$ выходная последовательность будет неограниченно возрастающей.

Разностное уравнение: $y(nT) = x(nT) + K \cdot y(nT - T)$.

Применим Z-преобразование: $Y(z) = X(z) + KY(z)z^{-1}$.

$$\text{Системная функция: } H(z) = \frac{Y(z)}{X(z)} = \frac{1}{1 - Kz^{-1}} = \frac{z}{z - K}.$$

Устойчивость: полюса функции должны лежать внутри единичной окружности в z -плоскости. В данном случае единственный полюс $z_1 = K$. Следовательно, система устойчива, если $|K| < 1$.

ИХ: $h(nT) = K^n$, $n \geq 0$, $|K| < 1$. На рис. 1.34 показаны графики ИХ при $K_1 = 0,9$; $K_2 = 0,5$; $K_3 = 0,25$.

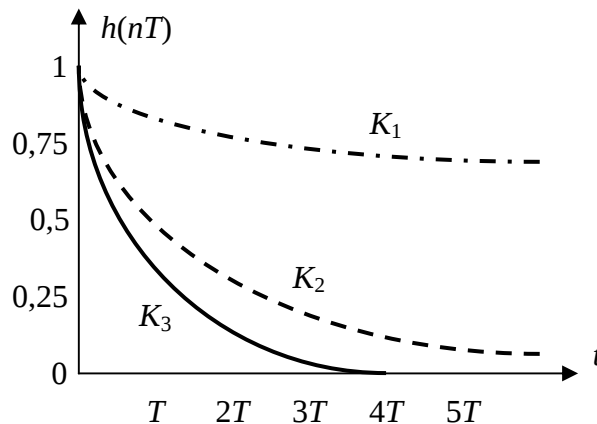


Рис. 1.34. – Графики импульсной характеристики СФ

Недостаток реализации СФ: при изменении времени сглаживания изменяется коэффициент передачи фильтра по постоянной составляющей (или изменяется коэффициент усиления фильтра).

$$A(\omega)|_{\omega=0} = |H(z)|_{z=e^{j\omega T}=1} = \frac{1}{1-Kz^{-1}} \Big|_{z=1} = \frac{1}{1-K}.$$

Чем выше степень сглаживания ($K \rightarrow 1$), тем больше коэффициент передачи по постоянной составляющей ($A(0) \rightarrow \infty$).

Используя выражение для прямого Z-преобразования, получим

$H(1) = \sum_{n=0}^{\infty} h(nT)$, т.е. сумма отсчетов импульсной характеристики равна коэффициенту передачи по постоянной составляющей.

Универсальная базовая ячейка (УБЯ) цифрового рекурсивного фильтра (ЦРФ) (рис. 1.35)

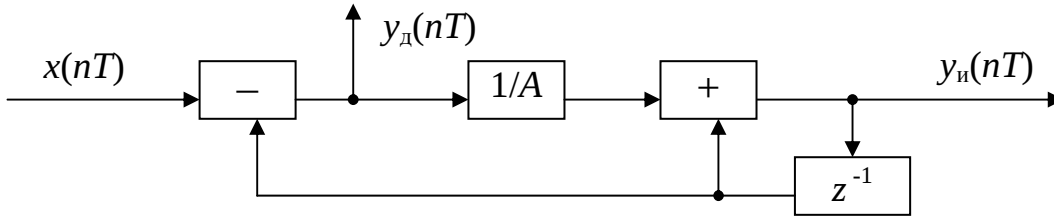


Рис. 1.35. – Структурная схема УБЯ

На рис. 1.35 $y_d(nT)$ – дифференцирующий выход; $y_u(nT)$ – интегрирующий выход; A – параметр, определяющий постоянную интегрирования или дифференцирования.

Таким образом, УБЯ выполняет функции интегрирования и дифференцирования и обладает постоянным коэффициентом передачи на $\omega = 0$.

Режим интегрирования (рис. 1.36):

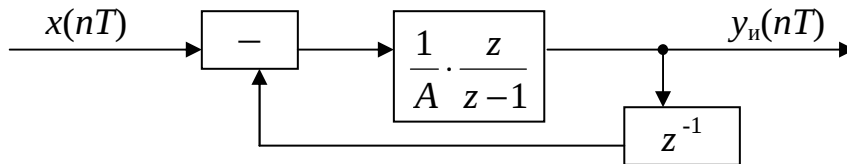


Рис. 1.36. – УБЯ в режиме интегрирования

$H(z) = \frac{z}{z-1}$ – это рассмотренный ранее цифровой интегратор.

$$\text{Системная функция: } H(z) = \frac{\frac{1}{A} \cdot \frac{z}{z-1}}{1 + \frac{1}{A} \cdot \frac{z}{z-1} z^{-1}} = \frac{\frac{1}{A} z}{z - \left(1 - \frac{1}{A}\right)} = \frac{Y_u(z)}{X(z)}.$$

$$\text{РУ: } y_u(nT) = \frac{1}{A} x(nT) + \left(1 - \frac{1}{A}\right) y(nT - T), n \geq 0.$$

Определим ИХ ЦФ при различных A .

$h_u(nT) = y_u(nT) \Big|_{x(nT)=\delta(nT)}$, при этом $y_u(nT) \equiv 0, n < 0$. Подставим значения $A_1 = 2, A_2 = 4, A_3 = 8$ в разностное уравнение и получим следующие результаты (табл. 1.4 и рис. 1.37):

Таблица 1.4

	$t = 0$	$t = T$	$t = 2T$	$t = 3T$	$t = 4T$	$t = 5T$
$A = 2$	0,5	0,25	0,125	0,0625	0,03	0,02
$A = 4$	0,25	0,19	0,14	0,11	0,08	0,06
$A = 8$	0,125	0,11	0,1	0,08	0,07	0,06

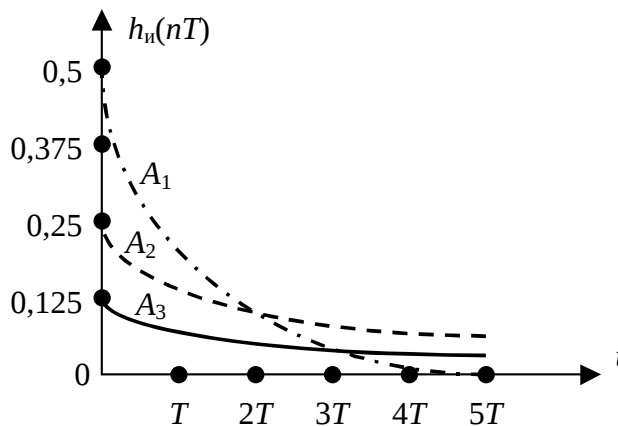


Рис. 1.37. – Импульсные характеристики УБЯ в режиме интегрирования

При любых значениях A сумма отсчетов ИХ равна 1. Это означает, что коэффициент передачи по постоянной составляющей сигнала равен 1.

$$H_u(z) \Big|_{z=1} = \sum_{n=0}^{\infty} h_u(nT) = \frac{1/A}{1 - (1 - 1/A)} = 1, \text{ т.е. } \sum_{n=0}^{\infty} h_u(nT) = 1 \text{ при любых } A.$$

Режим дифференцирования (рис. 1.38):

В данном случае системная функция имеет вид:

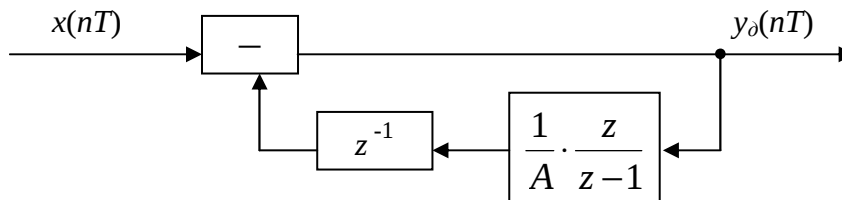


Рис. 1.38. – Математическая модель УБЯ в режиме дифференцирования

$$H_o(z) = \frac{Y_o(z)}{X(z)} = \frac{1}{1 + \frac{1}{A} \cdot \frac{z}{z-1} z^{-1}} = (1 - z^{-1}) \frac{1}{1 - \left(1 - \frac{1}{A}\right) z^{-1}}, \text{ как видно из вы-}$$

ражения, такая системная функция может быть получена последовательным

включением двух звеньев – вычислителя первой разности и сглаживающего фильтра (см. выше).

$$\text{РУ: } Y_{\partial}(z) = X(z) - X(z)z^{-1} + \left(1 - \frac{1}{A}\right)Y_{\partial}(z)z^{-1};$$

$$y_{\partial}(nT) = x(nT) - x(nT - T) + \left(1 - \frac{1}{A}\right)y_{\partial}(nT - T), \quad n \geq 0.$$

Найдем ИХ так же, как в предыдущем пункте для $A_1 = 2$, $A_2 = 4$, $A_3 = 8$ (табл. 1.5 и рис. 1.39):

Таблица 1.5

	$t = 0$	$t = T$	$t = 2T$	$t = 3T$	$t = 4T$	$t = 5T$
$A = 2$	1	-0,5	-0,25	-0,125	-0,063	-0,03
$A = 4$	1	-0,25	-0,19	-0,14	-0,11	-0,08
$A = 8$	1	-0,125	-0,11	-0,1	-0,08	-0,07

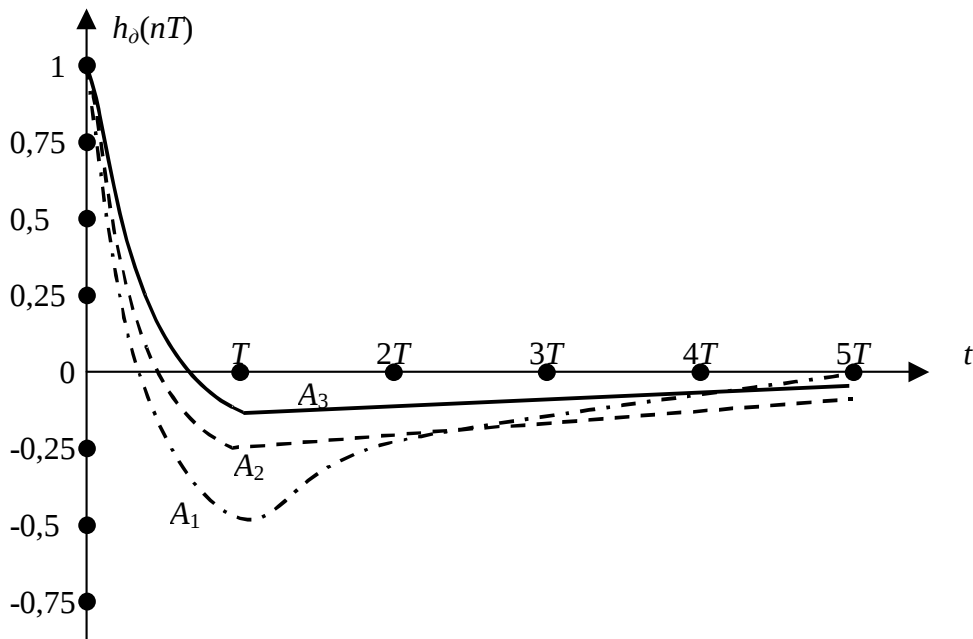


Рис. 1.39. – Импульсные характеристики УБЯ в режиме дифференцирования

Главный признак дифференцирующих цепей – они не пропускают постоянную составляющую сигнала. Это означает, что сумма отсчетов ИХ равна нулю:

$$H_{\partial}(1) = \frac{1-1}{1-(1-1/A)} = 0.$$

1.7. Нелинейные эффекты в цифровых фильтрах

1.7.1. Эффекты квантования

Рассмотрим цифровой рекурсивный фильтр первого порядка (рис. 1.40):

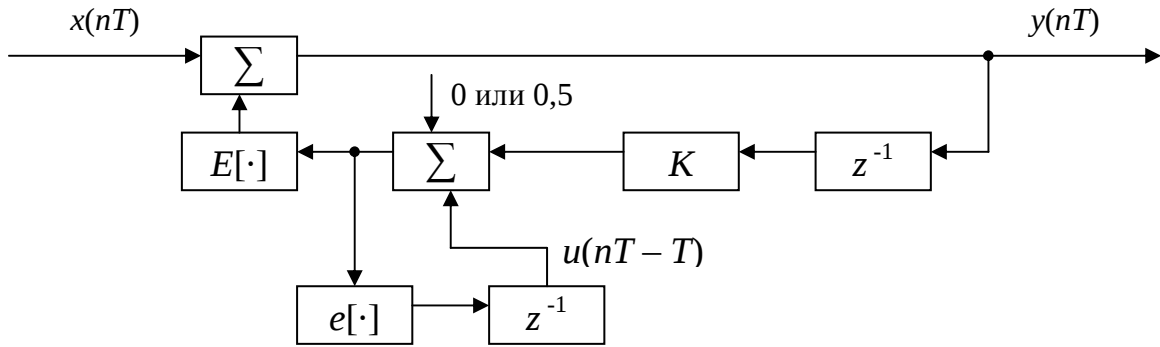


Рис. 1.40. – Цифровой рекурсивный фильтр с оператором квантования данных

Здесь $E[\cdot]$ – оператор взятия целой части числа; $e[\cdot]$ – оператор взятия дробной части числа (выделение остатка $u(nT)$). Добавление 0 или 0,5 определяет тип нелинейной операции – усечение либо округление отсчетов сигнала, соответственно, от этого зависит остаток $e[x] = x - E[x]$.

Ниже в табл. 1.6 приведены отсчеты выходных значений фильтров при подаче на их вход воздействий в виде однократного импульса с амплитудой десять единиц при $K=0,9$. При этом использованы следующие реализации рекурсивного фильтра.

1. Дискретный фильтр (без операции квантования данных по амплитуде). Его РУ: $y(nT) = x(nT) + Ky(nT - T)$, т.е. сглаживающий фильтр.

2. ЦФ с усечением данных (усечение – добавление нуля по схеме) описывается РУ: $y(nT) = x(nT) + E[Ky(nT - T)]$, $n \geq 0$.

3. ЦФ с учетом округления данных (добавление 0,5), его РУ имеет вид: $y(nT) = x(nT) + E[Ky(nT - T) + 0,5]$, $n \geq 0$.

4. ЦФ с учетом остатков от выполнения арифметических операций (АО):

$$y(nT) = x(nT) + E \left[Ky(nT - T) + \begin{Bmatrix} 0 \\ 0,5 \end{Bmatrix} + u(nT - T) \right], n \geq 0,$$

где $u(nT) = e \left[Ky(nT - T) + \begin{Bmatrix} 0 \\ 0,5 \end{Bmatrix} + u(nT - T) \right]$ – РУ остатка.

Таблица 1.6

Отклики ЦФ на входной сигнал $x(nT) = 10\delta(nT)$, $K = 0,9$

Тип фильтра	Период обработки данных												Примечания
	0	T	$2T$	$3T$	$4T$	$5T$	$6T$	$7T$	$8T$	$9T$	$10T$	$11T$	
1.	10	9	8,1	7,29	6,56	5,9	5,31	4,78	4,30	3,87	3,49	3,24	
	10	9	8	7	7	6	5	5	4	4	3	3	Округление по выходу
2.	10	9	8	7	6	5	4	3	2	1	0	0	
3.	10	9	8	7	6	5	5	5	5	5	5	5	
4а. (Усечение)	10	9	8	7	6	6	5	4	4	4	3	3	
	0	0	0,1	0,3	0,6	0	0,4	0,9	0,5	0,1	0,7	0,4	$u(nT)$
4б. (Округление)	10	9	8	7	7	6	5	5	4	4	3	3	
	0	0	0,1	0,3	-0,4	-0,1	0,3	-0,2	0,3	-0,1	0,5	0,2	$u(nT)$

Анализируя отклики фильтра на однократное воздействие, можно сделать вывод о том, что наиболее точным методом рекурсивной обработки цифровых данных является метод выполнения арифметических операций с сохранением и использованием при обработке дробных частей результатов. Фильтр с усечением данных укорачивает импульсную характеристику фильтра, а при округлении данных в фильтре возникают паразитные колебания, связанные с предельными циклами, а импульсная характеристика в данном случае существенно удлиняется. В фильтре с учетом дробных частей результатов вычислений вид выходного сигнала наиболее близок к отклику точного (дискретного) фильтра (рис. 1.41).

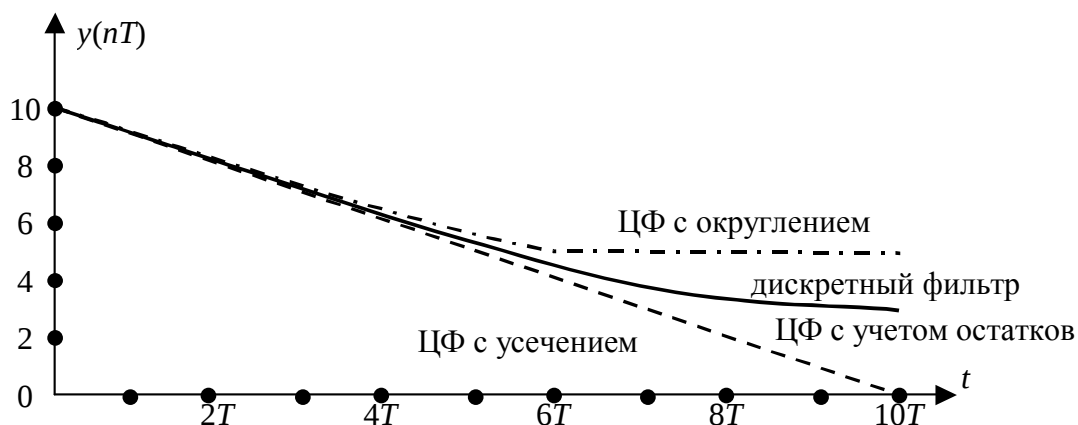


Рис. 1.41. – Отклики ЦФ с различными операциями квантования

Предельный цикл:

При некоторых значениях коэффициентов обратной связи (ОС) из-за нелинейности операций квантования в ЦФ возникает неустойчивый режим, называемый *предельным циклом*. Проявляется предельный цикл в виде колебаний, либо появлением «мертвой» зоны. Условием возникновения этого режима является расположение полюсов фильтра вблизи единичной окружности. В рассматриваемом примере предельный цикл возникает при $0,5 < |K| < 1$ (K – коэффициент ОС, здесь $K = b_1$).

1.7.2. Ошибки квантования в ЦРФ

Рассмотрим ошибки вычислений из-за ограничения разрядности. Запишем разностное уравнение с учетом операций квантования:

$$y(n) = \sum_{i=0}^M R[a_i x(n-i)] + \sum_{j=1}^N R[b_j y(n-j)], \quad n = 0, 1, 2, \dots,$$

где $R[\cdot]$ – оператор квантования, т.е. (для j -го компонента выходной последовательности) (рис. 1.42):

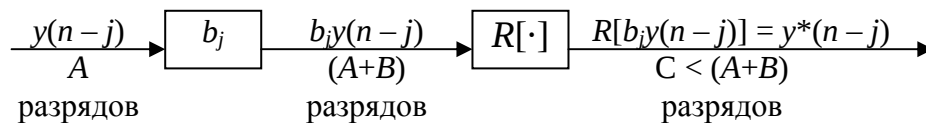


Рис. 1.42. – Ограничение разрядности выходного результата

Здесь C – разрядность арифметического устройства по выходу.

Аналогично для $(j+1)$ -го компонента (такта) получим $C < (A+B)+B$ и т.д., т.е. происходит лавинообразное увеличение разрядности выходного результата. Видно, что из-за конечного значения C ошибка выходного сигнала со временем увеличивается.

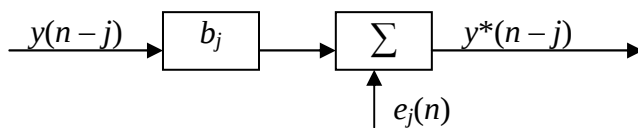


Рис. 1.43. – Линеаризация цепи ОС

При вероятностной оценке ошибки квантования операцию квантования линеаризируем, т.е. вводим сумматор и источник шума $e_j(n)$. При этом возникают шумы кван-

тования с дисперсией (рис. 1.43): $\sigma_j^2 = \frac{\Delta_j^2}{12}$, $\Delta_j = 2^{-C_j}$. C_j – номер младшего разряда в j -й цепи. Задавая конкретно разрядность C дробной части, мы теряем информацию Δ_j .

Рассмотрим, как ведут себя ошибки округления в цифровом рекурсивном фильтре первого порядка при различных формах реализации (прямая и каноническая). Системная функция такого фильтра в общем случае имеет вид:

$H(z) = \frac{1 + a_1 z^{-1}}{1 - b_1 z^{-1}}$. Линеаризуя схему, ставим сумматоры.

Локальная системная функция: $H_{1,2}(z) = \frac{1}{1 - b_1 z^{-1}}$.

Локальная ИХ: $h_{1,2}(n) = b_1^n$. Интервал дискретизации T опустим, т.к. при цифровой реализации он не существен, особенно при обработке на ЭВМ (он даже нам зачастую неизвестен).

$$\sigma_{k \text{ вых}}^2 = \sigma_{k \text{ вх}}^2 \sum_{n=0}^{\infty} h_k^2(n).$$

Реализуем ЦРФ, используя прямую форму реализации (рис. 1.44).

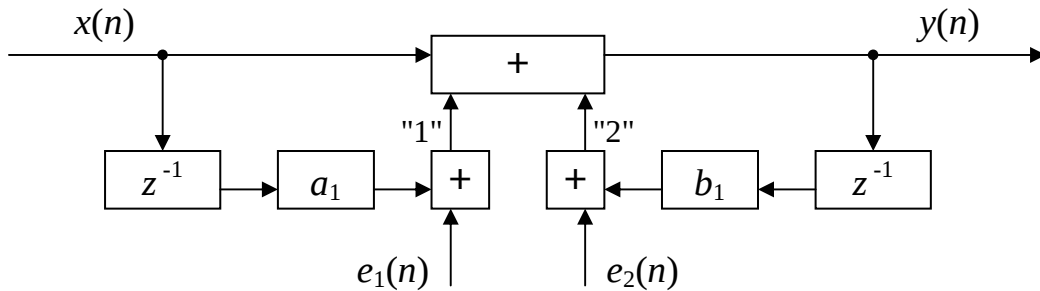


Рис. 1.44. – Прямая форма реализации ЦРФ

Поясним рис. 1.44: при вероятностной оценке ошибок мы имеем право с известной точностью заменить операцию ограничения разрядности результата введением дополнительных источников шума $e_1(n)$ и $e_2(n)$.

Ошибки $e_1(n)$ и $e_2(n)$ приложены к одной точке. Пусть $\Delta_1 = \Delta_2 = \Delta$, тогда $\sigma_{\Sigma \text{ вх}}^2 = \sigma_1^2 + \sigma_2^2 = \frac{\Delta_1^2}{12} + \frac{\Delta_2^2}{12} = \frac{\Delta^2}{6}$, а $\sigma_{\Sigma \text{ вых}}^2 = \sigma_{\Sigma \text{ вх}}^2 \sum_{n=0}^{\infty} h_{1,2}^2(n) = \frac{\Delta^2}{12} \cdot \frac{2}{1 - b_1^2}$ (это следует из формулы суммы членов бесконечной геометрической прогрессии).

Отметим, что при синтезе ЦФ недостаточно спроектировать структуру фильтра, необходимо также, что бы ошибки округления входили с минимальными коэффициентами. В общем случае это нетривиальная задача.

При канонической форме реализации ЦРФ его структурная схема представлена на рис. 1.45.

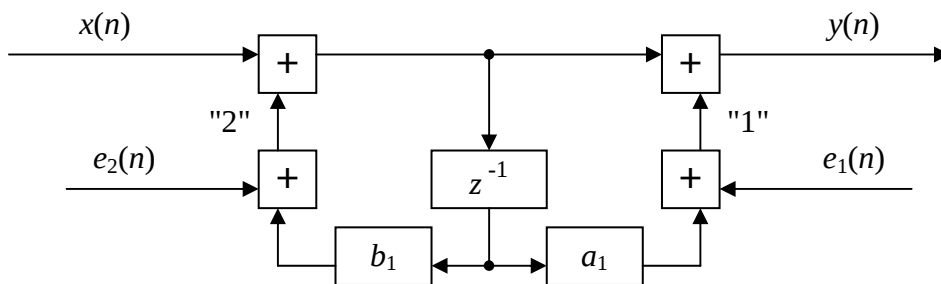


Рис. 1.45. – Каноническая форма реализации ЦРФ

Найдем дисперсию:

$$\sigma_{1\text{блх}}^2 = \frac{\Delta_1^2}{12}; \quad \sigma_{2\text{блх}}^2 = \sigma_{2\text{бх}}^2 \frac{1}{2\pi j} \oint_C H_2(z) H_2(z^{-1}) z^{-1} dz = \sigma_{2\text{бх}}^2 \sum_{k=1}^K \operatorname{Res}_{z=z_k} \Psi(z).$$

$$H_2(z) = H(z) = \frac{z + a_1}{z - b_1}, \quad \Psi(z) = \frac{z + a_1}{z - b_1} \cdot \frac{1 + a_1 z}{1 - b_1 z} \cdot \frac{1}{z}.$$

$$\text{Есть три полюса: } z_1 = b_1; z_2 = 1/b_1; z_3 = 0, \text{ т.е. } \sigma_{2\text{блх}}^2 = \sigma_{2\text{бх}}^2 \frac{1 + 2a_1 b_1 + a_1^2}{1 - b_1^2}.$$

Учитываем полюса $H(z)$ и $1/z$ и не учитываем $H(z^{-1})$. Пусть для шумов e_1 и e_2 выполняется равенство $\Delta_1 = \Delta_2 = \Delta$. Тогда:

$$\sigma_{\Sigma \text{блх}}^2 = \sigma_{1\text{блх}}^2 + \sigma_{2\text{блх}}^2 = \frac{\Delta^2}{12} + \frac{\Delta^2}{12} \cdot \frac{1 + 2a_1 b_1 + a_1^2}{1 - b_1^2}. \text{ При этом, если знаки } a_1 \text{ и } b_1$$

разные, то $\sigma_{2\text{блх}}^2$ уменьшается!

Пример 1: Пусть $a_1 = -0,9$ и $b_1 = 0,9$. Тогда для прямой формы реализации $\sigma_{\Sigma \text{блх}}^2 = \frac{\Delta^2}{12} \cdot \frac{2}{1 - b_1^2} = \frac{\Delta^2}{12} 10,526$, а при канонической форме реализации вы-

ражение для дисперсии будет иметь вид $\sigma_{\Sigma \text{блх}}^2 = \frac{\Delta^2}{12} 2$, т.е. выигрыш в 5,26 раза.

Пример 2: Пусть $a_1 = 0,9$ и $b_1 = 0,9$. Тогда для прямой формы реализации $\sigma_{\Sigma \text{блх}}^2 = \frac{\Delta^2}{12} 10,526$, а для канонической формы реализации $\sigma_{\Sigma \text{блх}1}^2 = \frac{\Delta^2}{12} 19,052$, т.е. проигрыш в 1,81 раза по сравнению с прямой формой реализации ЦФ.

Таким образом, при разнополярных коэффициентах a_1 и b_1 лучше каноническая форма реализации, а при однополярных – прямая.

1.7.3. Методы борьбы с нелинейными эффектами в рекурсивных ЦФ

Как было показано выше, численные методы, используемые при цифровых вычислениях, дают лишь приближенные результаты из-за конечной разрядности арифметических устройств. Особенно сильно сказывается влияние ошибок квантования результатов арифметических операций, которое проявляется при выполнении рекурсивных вычислений. Это обстоятельство существенно ограничивает область применения рекурсивных алгоритмов при их реализации в цифровых фильтрах с пониженной разрядностью арифметических устройств. В общем виде разностное уравнение РЦФ с учетом операций квантования записывается следующим образом:

$$y(n) = \sum_{i=0}^M R[a_i x(n-i)] + \sum_{j=1}^N R[b_j y(n-j)], \quad n \geq 0,$$

где $x(n)$ – входные данные; $y(n)$ – выходные данные; a_i и b_j – коэффициенты фильтрации (для нерекурсивной и рекурсивной частей фильтра, соответственно); $R[\cdot]$ – оператор квантования; n – порядковый номер данных (период дискретизации T в данном случае принят равным единице).

Квантование результатов арифметических операций при вычислениях приводит к возникновению предельных циклов и, как следствие, к снижению точности вычислений.

Известно несколько способов борьбы с предельными циклами при рекурсивных вычислениях:

1. Добавление случайных чисел (шума) к каждому из членов квантованной суммы – это приводит к срыву колебаний предельного цикла, но ошибка вычислений увеличивается.

2. Модификация алгоритмов вычислений или поиск оптимальной формы реализации ЦФ. При модификации алгоритма вычислений (в сторону его усложнения) иногда удается устранить предельные циклы (или уменьшить их влияние), но это удастся лишь для нескольких частных случаев рекурсивных фильтров и, поэтому, данный метод не универсален.

3. Наиболее универсальным и точным методом рекурсивных вычислений является способ подавления предельных циклов с одновременным увеличением точности фильтрации за счет использования остатков от выполнения арифметических операций при последующих вычислениях.

Разностное уравнение для линейных рекурсивных ЦФ произвольного порядка с учетом остатков от выполнения арифметических операций (умножения, например) имеет вид:

$$y(n) = \sum_{i=0}^M R[a_i x(n-i) + u_i(n-1)] + \sum_{j=1}^N R[b_j y(n-j) + u_j(n-1)], \quad n \geq 0,$$

где $u_i(n)$ и $u_j(n)$ – остатки от выполнения арифметических операций, причем

$$u_i(n) = e[a_i x(n-i) + u_i(n-1)];$$

$$u_j(n) = e[b_j y(n-j) + u_j(n-1)];$$

$$e[x] = x - R[x].$$

Здесь $R[\cdot]$ – оператор квантования (определение целой части); $e[\cdot]$ – операция нахождения остатка (дробной части числа).

Следует отметить, что учитывать остатки от выполнения арифметических операций можно не во всех «ветвях» алгоритма ЦФ, а лишь в «критических» (максимально ухудшающих точность обработки), начиная с рекурсивных «ветвей».

Выражение для вычисления, например, $u_j(n)$ будет таким:

$$u_j(n) = b_j y(n-j-1) - R[b_j y(n-j-1) + u_j(n-1)] + u_j(n-1), \quad n \geq 0.$$

Введение дополнительных цепей коррекции (для учета значений остатков) существенно не изменяет передаточную функцию, но приводит к декорреляции ошибок округления, что уменьшает их влияние на результат вычислений.

Таким образом, указанный метод обладает повышенной точностью вычислений по сравнению с другими методами. Дополнительным преимуществом использования *метода остатков* является то, что при выполнении арифметических операций их остаток, как правило, не приходится дополнительно вычислять, т.е. повышается быстродействие вычислений при заданной точности, и есть возможность реализации прецизионных рекурсивных вычислений без существенного усложнения алгоритма, реализуемого в арифметическом устройстве ЦФ.

Рассмотрим математическую модель умножителя с сохранением остатка (рис. 1.46).

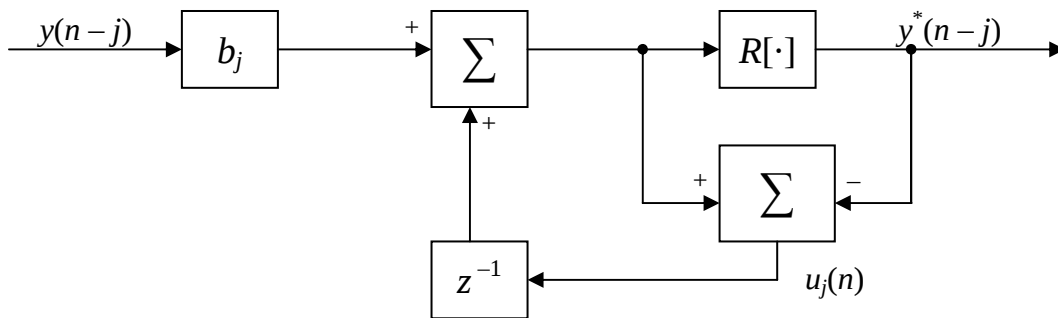


Рис. 1.46. – Математическая модель умножителя с сохранением остатка (произвольная j -я ветвь ОС)

Здесь $u_j(n)$ – сигнал коррекции (остаток);

$y^*(n-j) = R[b_j y(n-j) + u_j(n-1)]$ – выходной сигнал;

$u_j(n) = b_j y(n-j) - y^*(n-j)$, $n \geq 0$, $T = 1$ с.

Используя вероятностную оценку ошибок линеаризуем j -ю цепочку ОС. Сначала представим j -ю цепочку ОС в виде (рис. 1.47):

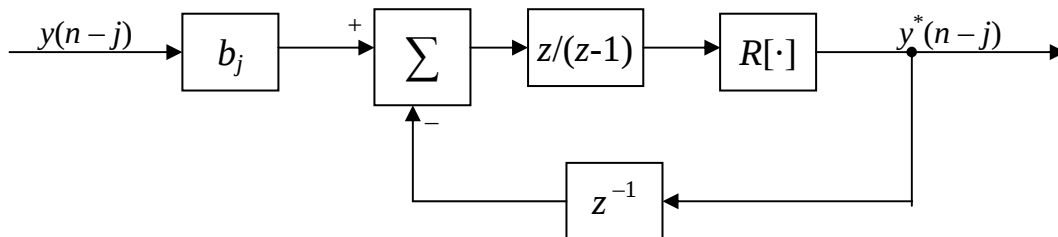


Рис. 1.47. – Преобразованная схема умножителя с сохранением остатка

Затем линеаризуем эту модель (рис. 1.48):

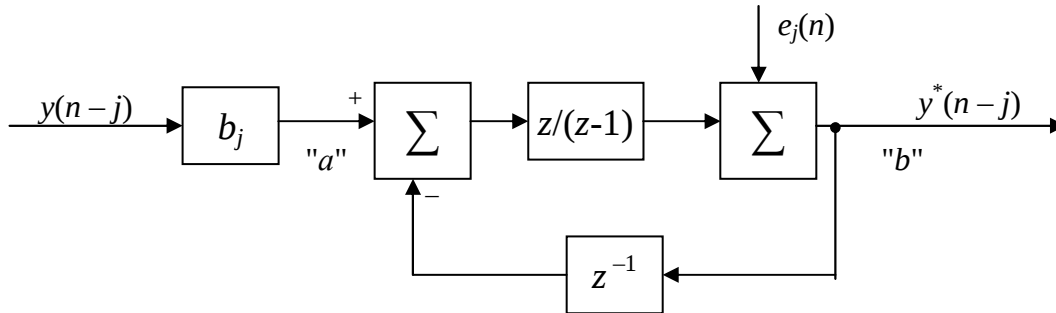


Рис. 1.48. – Линеаризованная схема умножителя с сохранением остатка

Системная функция передачи по сигналу от a до b :

$$H_s(z) = \frac{\frac{z}{z-1}}{1 + \frac{z}{z-1}z^{-1}} = \frac{z}{z-1+1} \equiv 1.$$

Системная функция передачи по шуму в точке b :

$$H_l(z) = \frac{1}{1 + \frac{z}{z-1}z^{-1}} = \frac{z-1}{z} = 1 - z^{-1} \text{ – вычислитель первой разности (циф-}$$

ровой дифференциатор).

С учетом этого преобразуем схему (рис. 1.49):

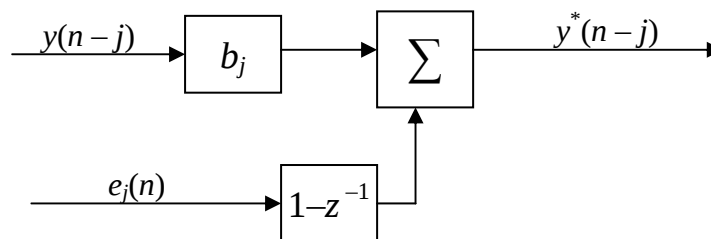


Рис. 1.49. – Эквивалентная схема умножителя с сохранением остатка

Получается, что при учете остатков арифметических операций шум дополнительно проходит через цифровой дифференциатор.

1.7.4. Точность и эффективность цифровых вычислений с сохранением остатков

Произведем оценку эффективности цифровых вычислений с использованием остатков на примере рекурсивного фильтра второго порядка (ЦРФ2П) (рис. 1.50).

Разностное уравнение без сохранения остатков:

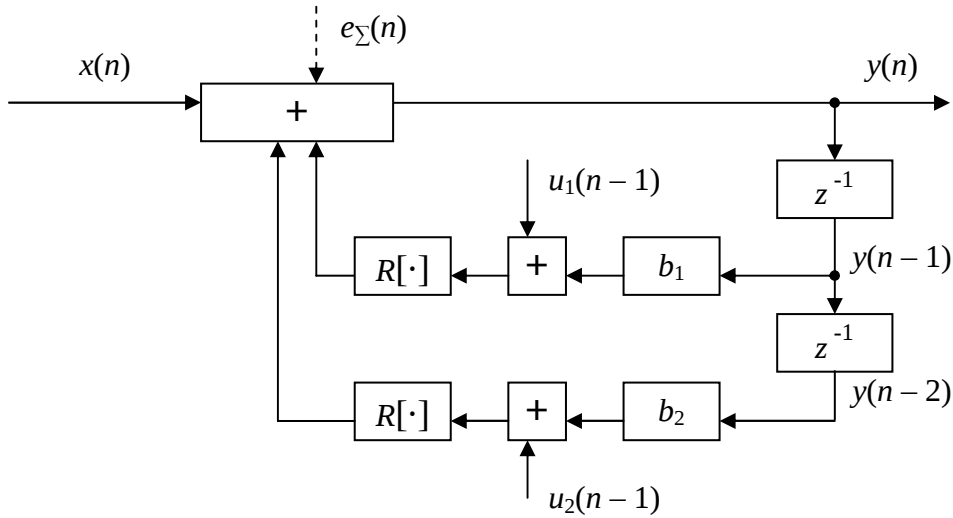


Рис. 1.50. – Структурная схема ЦРФ2П

$$y_1(n) = x(n) + R[b_1 y_1(n-1)] + R[b_2 y_1(n-2)], \quad n \geq 0.$$

Разностное уравнение с сохранением остатков:

$$y_2(n) = x(n) + R[b_1 y_2(n-1) + u_1(n-1)] + R[b_2 y_2(n-2) + u_2(n-1)], \quad n \geq 0,$$

$$u_1(n) = b_1 y_2(n-1) + u_1(n-1) - R[b_1 y_2(n-1) + u_1(n-1)],$$

$$u_2(n) = b_2 y_2(n-2) + u_2(n-1) - R[b_2 y_2(n-2) + u_2(n-1)].$$

Начальные условия: $T = 1$ с; $y_2(-1) = y_2(-2) = 0$; $u_1(-1) = u_2(-1) = 0$.

Найдем дисперсию шума округления. Т.к. шумы приложены к одной точке, а именно, ко входу фильтра, то дисперсия ошибки вычислений на выходе фильтра:

$$\sigma_{s1}^2 = \frac{\Delta^2}{6} \sum_{n=0}^{\infty} h^2(n) = \frac{\Delta^2}{6} \cdot \frac{1-b_2}{(1+b_2)[(1-b_2)^2 - b_1^2]}.$$

Дисперсия ошибки

при сохранении остатков:
$$\sigma_{s2}^2 = \frac{\Delta^2}{6} \cdot \frac{2(1-|b_1|-b_2)}{(1+b_2)[(1-b_2)^2 - b_1^2]}.$$

Оценим эффективность предлагаемого метода с помощью отношения

дисперсий:
$$D = \frac{\sigma_{s2}^2}{\sigma_{s1}^2} = \frac{2(1-|b_1|-b_2)}{1-b_2}.$$
 Представим графически дисперсии оши-

бок σ_{s1}^2 и σ_{s2}^2 (рис. 1.52 и 1.53) в виде полей на «треугольнике устойчивости» (рис. 1.51). На рис. 1.52, 1.53 плоские части соответствуют усилению ошибок.

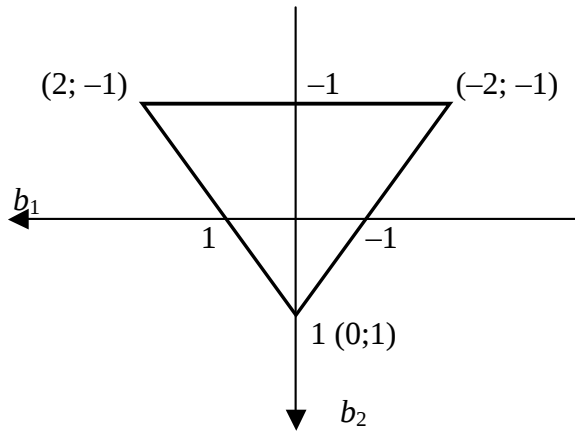


Рис. 1.51. – Треугольник устойчивости ЦРФ2П

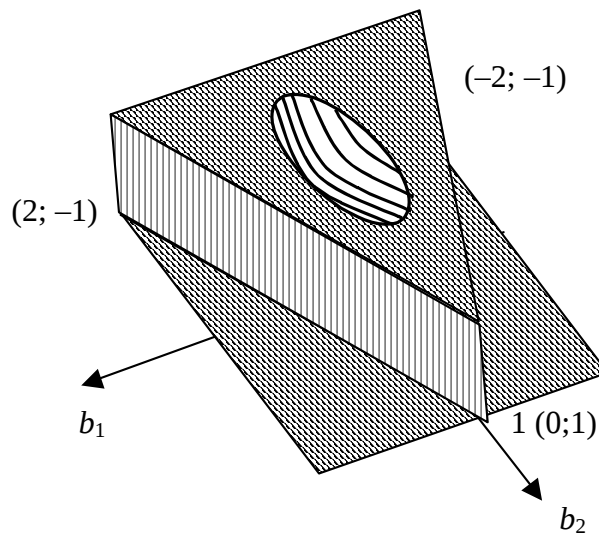


Рис. 1.52. – $\sigma_{s1}^2(b_1; b_2)$ (ограничение по уровню $\Delta^2/6$)

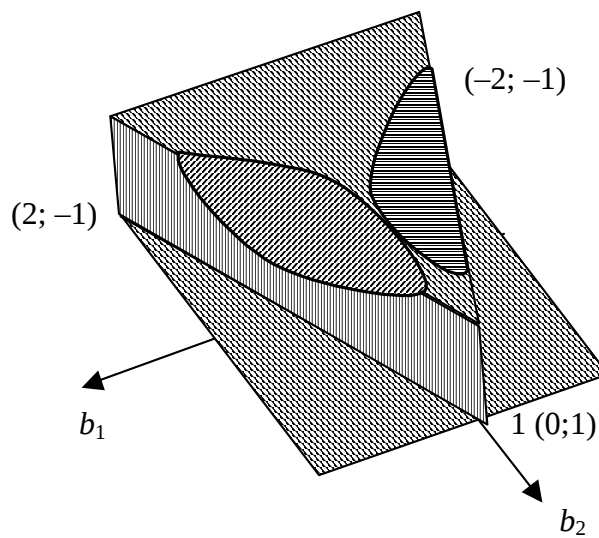


Рис. 1.53. – $\sigma_{s2}^2(b_1; b_2)$ (ограничение по уровню $\Delta^2/6$)

Отношение дисперсий: $D = \frac{\sigma_{s2}^2}{\sigma_{s1}^2}$ (рис. 1.54)

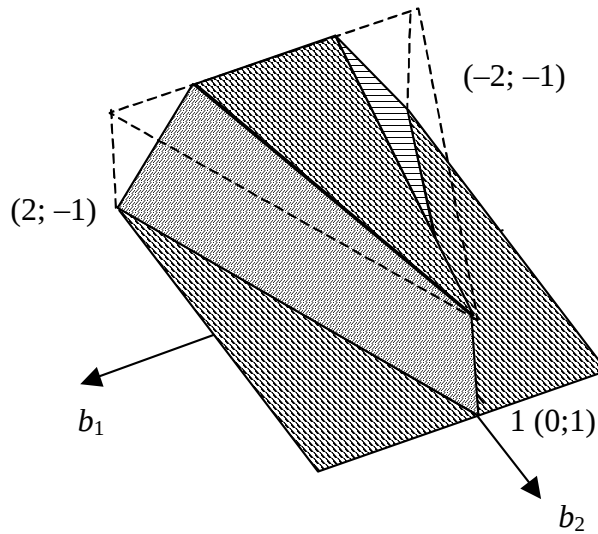


Рис. 1.54. – D (ограничение по уровню 1)

Выигрыш происходит на наклонных участках $|b_1| > 0,5$. В этих областях предпочтительнее фильтр с использованием остатков, а на плоских областях лучше использовать фильтр без учета остатков.

Оценим точность фильтрации данных в ЦРФ на основе УБЯ при традиционном алгоритме умножения данных и при умножении данных с сохранением остатков (рис. 1.55).

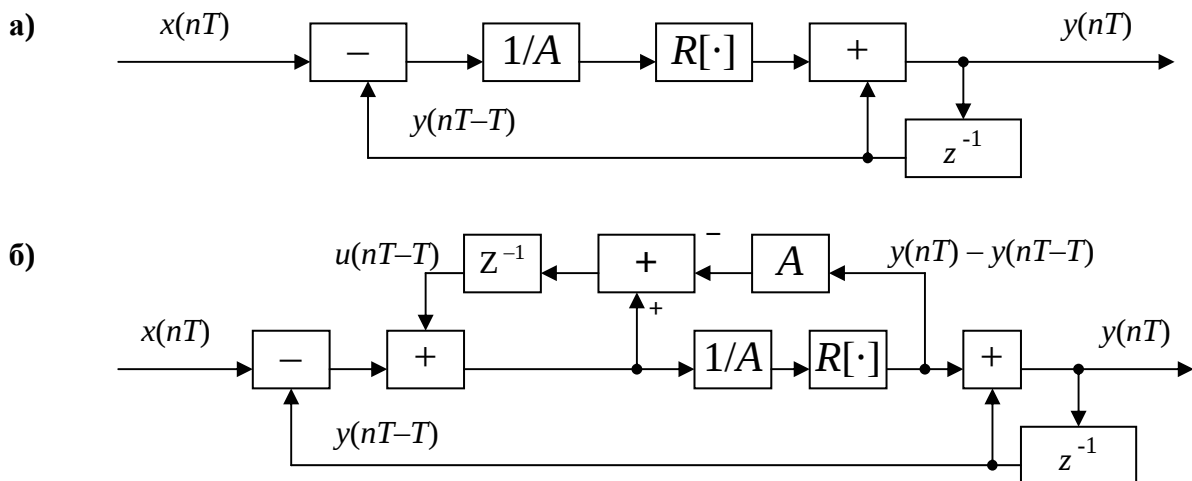


Рис. 1.55. – Структурная схема ЦРФ на основе УБЯ:

а) операция умножения с квантованием данных;

б) операция умножения с сохранением остатка.

Разностное уравнение:

$$y_1(nT) = R\{[x(nT) - y_1(nT - T)]/A\} + y_1(nT - T), \quad n \geq 0 \text{ – без учета остатков.}$$

$y_2(nT) = R\{[x(nT) - y_2(nT - T) + u(nT - T)]/A\} + y_2(nT - T)$, $n \geq 0$ – с учетом остатков.

Если убрать $u(nT - T)$, то получится первое РУ. Для формирования сигнала остатка РУ имеет вид:

$u(nT) = x(nT) - y_2(nT - T) + u(nT - T) - A[y_2(nT) - y_2(nT - T)]$ – это остаток от деления на предыдущем шаге.

Дисперсия ошибки округления результата на выходе ЦФ:

1. $\sigma_1^2 = \frac{\Delta^2}{12} \cdot \frac{A^2}{2A-1}$ – фильтр без учета остатков;

2. $\sigma_2^2 = \frac{\Delta^2}{12} \cdot \frac{2A}{2A-1}$ – с учетом остатков.

При $A \gg 1$ сильное сглаживание: $\sigma_1^2 \cong \frac{\Delta^2}{12} \cdot \frac{A}{2}$; $\sigma_2^2 \cong \frac{\Delta^2}{12}$.

На рис. 1.56 приведены следующие зависимости:

- 1 – расчетная зависимость для обычного фильтра;
- 2 – расчетная зависимость для фильтра с учетом остатков;
- (далее экспериментальные зависимости)
- 3 – обычный фильтр;
- 4 – с учетом остатков;
- 5 – фильтр с добавлением шума;
- 6 – фильтр с управляемым округлением.

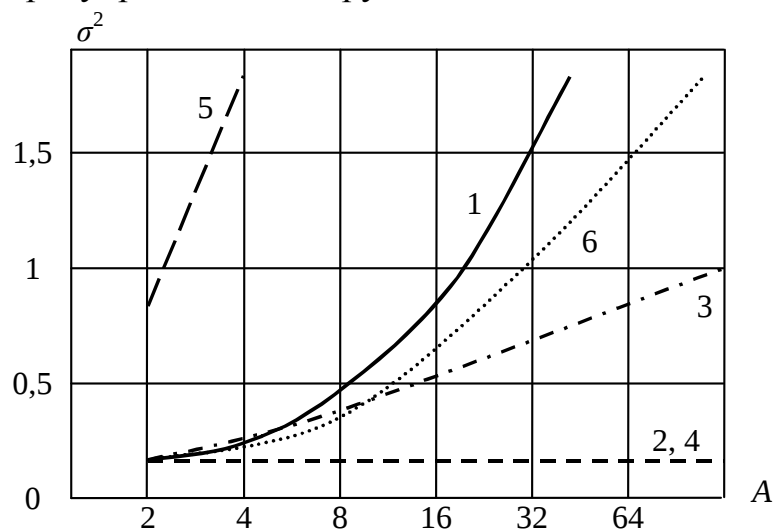


Рис. 1.56. – Расчетные зависимости шумов квантования и результаты вычислительного эксперимента ($\Delta = 1$)

Таким образом, в ЦРФ по методу округления с учетом остатков ошибки округления практически не зависят от параметра фильтрации A , а теоретические данные хорошо совпадают с экспериментальными. Дисперсия ошибки для алгоритма без учета остатков при малых A несколько больше, а для больших A меньше расчетной. Дисперсия ошибки, вносимой в вычислениях, при использовании алгоритма с управляемым округлением, близка к расчетной (для алгоритма без учета остатков при $A \leq 16$). Наибольшая дисперсия ошибки вычислений получается для алгоритма с добавлением случайных чисел перед округлением результата, так как для устранения предельных циклов в этом случае следует добавлять шум достаточной мощности.

Использование известных способов борьбы с предельными циклами при выполнении цифровых рекурсивных вычислений (введение случайных чисел небольшой мощности перед округлением, управляемое округление) не приводит к повышению точности вычислений, хотя предельные циклы в этих случаях устраняются. Алгоритм, учитывающий остатки от выполнения арифметических операций, обладает повышенной точностью вычислений. В частности, для рекурсивного алгоритма первого порядка с использованием остатков, происходит полное устранение предельных циклов и существенное уменьшение дисперсии ошибки вычислений (в $5 \div 10$ раз при значениях параметра фильтрации $A \leq 16$). Введение цепей учета остатков по методу остатков приводит к декорреляции ошибок округления, что уменьшает их влияние на результат вычислений. Следует отметить, что метод использования остатков согласуется с общим подходом по частотной коррекции шумов округления, применяемой для оптимизации точностных характеристик цифровых фильтров. Дополнительным преимуществом использования метода остатков является то, что при выполнении арифметических операций этот остаток, как правило, не приходится дополнительно вычислять. Остаток от деления получается естественным образом в регистрах после заданного числа итераций по вычислению искомого результата.

Данный метод коррекции ошибок округления применим и к нерекурсивным алгоритмам обработки данных, в особенности, когда эти результаты в дальнейшем используются при рекурсивных вычислениях как с фиксированной, так и с плавающей запятой. При реализации алгоритма фильтрации в вычислителе следует учитывать, что остатки от выполнения арифметических операций могут быть получены естественным образом в процессе вычислений. Это обстоятельство повышает быстродействие вычислений при заданной точности и дает возможность реализации прецизионных рекурсивных вычислений без существенного усложнения алгоритма, реализуемого в вычислителе.

1.7.5. Особенности построения каналов слежения с использованием ЦФ

В качестве примера практического использования метода остатков в цифровых контурах авторегулирования рассмотрим одно из устройств телевизионной автоматики.

При построении телевизионных систем, определяющих координаты подвижного объекта, как правило, используются два канала слежения по строке и кадру (X и Y), состоящих из *дискриминатора, траекторного фильтра и генератора стробов*. Такое построение следящих каналов устройства является общепринятым в телевизионной автоматике. В ТИС при измерении других параметров объекта, например, его яркости, также используются следящие каналы подобного типа. Следует отметить, что при цифровой реализации следящего канала возникают ошибки измерения координат объекта в виде смещения кода выходной координаты относительно истинной, а также возникают периодические автоколебания, амплитуда и период которых зависит от положения центра объекта относительно положения отсчетов видеосигнала по времени. Так при использовании в канале слежения одного цифрового интегратора частота колебаний пропорциональна относительному смещению энергетического центра объекта, а амплитуда колебаний в среднем равна 0,25 элемента разложения. Величину ошибок, возникающих в цифровом координаторе за счет нелинейных преобразований сигнала, можно в значительной степени уменьшить путем коррекции выходного сигнала дискриминатора дробной частью выходного сигнала траекторного фильтра, причем управление положением строба выполняется целой частью этого сигнала.

На рис. 1.57 приведена структурная схема строчного канала слежения с коррекцией сигнала рассогласования.

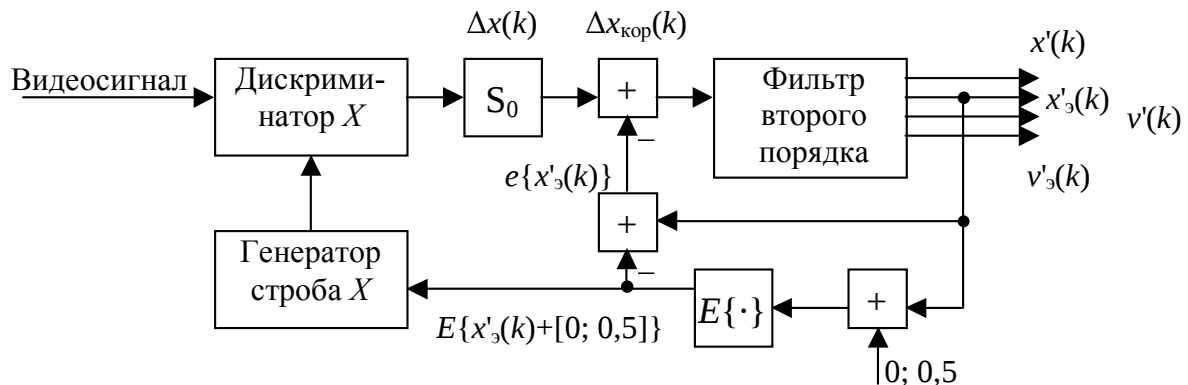


Рис. 1.57. – Структурная схема строчного канала слежения с коррекцией сигнала рассогласования

С выхода дискриминатора X снимается сигнал рассогласования центров строба и объекта по строке $\Delta x(k)$. Целая часть выходной координаты по X , которая управляет положением строчного строба, вычисляется блоком $E\{\cdot\}$ по значению выходной величины фильтра второго порядка $x'_y(k)$. Причем

выполняется либо усечение величины выходной координаты $E\{x'_3(k)\}$, либо ее округление $E\{x'_3(k) + 0,5\}$. Дробная часть выходной координаты $e\{x'_3(k)\}$ вычисляется как разность выходного значения фильтра и его квантованного аналога: $e\{x'_3(k)\} = x'_3(k) - E\left\{x'_3(k) + \begin{bmatrix} 0,5 \\ 0 \end{bmatrix}\right\}$. После вычисления корректированного значения сигнала рассогласования центров строба и объекта $\Delta x_{\text{кор}}(k) = S_0 \Delta x(k) - e\{x'_3(k)\}$ с помощью алгоритма траекторной фильтрации вычисляются оценки скорости и положения в текущем кадре и их экстраполяции на следующий кадр (рис. 1.58).

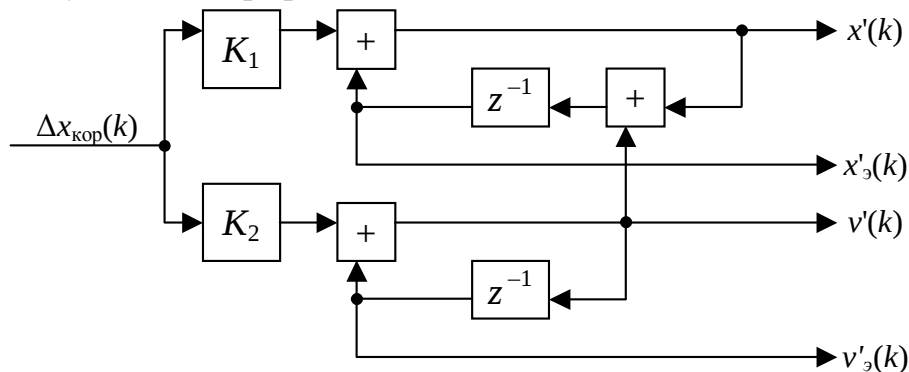


Рис. 1.58. – Математическая модель траекторного фильтра второго порядка

В случае использования дискриминатора с нормированием сигнала ошибки на сумму отсчетов сигнала от объекта и его вычисления с помощью пилообразной измерительной функции крутизна дискриминатора близка к единице лишь при отсутствии фоновой составляющей, по этому для исследования в структурную схему (рис. 1.57) введен блок, позволяющий изменять значение крутизны S_0 дискриминатора. Определим выражения, по которым вычисляются выходные сигналы фильтра второго порядка. Уточненные на k -шаге оценки положения и скорости, соответственно, равны:

$$x'(k) = x'_3(k) + K_1 \Delta x_{\text{кор}}(k),$$

$$v'(k) = v'_3(k) + K_2 \Delta x_{\text{кор}}(k).$$

Экстраполированные значения оценок положения и скорости объекта на следующий шаг, соответственно, равны

$$x'_3(k+1) = x'(k) + v(k), \quad v'_3(k+1) = v(k).$$

Структурная схема траекторного фильтра второго порядка, реализующего приведенные выше разностные уравнения, представлена на рис. 1.58.

Для устройства, структурная схема которого приведена на рис. 1.57, произведем оценку ошибки измерения координат в зависимости от значения крутизны S_0 и способа выделения целой и дробной частей выходного результата цифрового фильтра. На первый вход дискриминатора подается оцифрованный массив видеосигнала, а на его второй вход – опорный сигнал от гене-

ратора стробов, который определяет вид и положение измерительной функции дискриминатора. На выходе дискриминатора формируется сигнал рассогласования центров строба и объекта $\Delta x(k)$. Сигнал рассогласования $\Delta x(k)$ далее обрабатывается траекторным фильтром (рис. 1.58) и из его выходного результата выделяется целая $E\{x_3(k)\}$ и дробная $e\{x'_3(k)\}$ части результата. Дробная часть результата используется при вычислении скорректированного сигнала рассогласования $\Delta x_{кор}(k)$, а по целой части результата устанавливается положение строба в текущем кадре.

Для определения значений математического ожидания $m_{\Delta x}$ и среднеквадратичного отклонения $\sigma_{\Delta x}$ сигнала рассогласования центров объекта и строба $\Delta x(k)$ следует произвести усреднение данных по множеству кадров, причем параметром усреднения является пространственное положение объекта. Подобное усреднение следует реализовать и за счет движения объекта с постоянной скоростью («→» – слева–направо), («←» – справа–налево), причем значение скорости движения объекта необходимо задавать таким, чтобы его сдвиг за кадр был много меньше элемента разложения (например, $v_{об} = \pm 0,075$ эл./кадр). Для неподвижного объекта при $v_{об} = 0$ выполнялось усреднение $m_{\Delta x}$ и $\sigma_{\Delta x}$ по множеству кадров при сдвигах объекта «→» и «←» в пределах одного элемента разложения с шагом 0,075 эл.

На рис. 1.59 и 1.60 приведены результаты вычислительного эксперимента по измерению математического ожидания $m_{x'}$ и среднеквадратичного отклонения $\sigma_{x'}$ ошибки оценки координат подвижного объекта в зависимости от значений крутизны дискриминатора, то есть статистические характеристики разности истинной траектории его движения и ее оценки с помощью цифрового координатора. При *усечении* результата в цифровом следящем канале оценка траектории объекта смещена на 0,5 элемента и ее отклонение от истинной в среднем составляет 0,15 элемента. При *округлении* результата оценка траектории существенно не смещена, а $\sigma_{x'}$ также равна 0,15 элемента. При *учете дробной части* результата в канале слежения, как в случае усечения, так и округления, $\sigma_{x'}$ значительно меньше чем без учета дробной части, причем наибольший выигрыш (5...10 раз) реализуется при значениях крутизны $S_0 = 0,8...1$. Таким образом, при обеспечении значения крутизны дискриминатора близкого к 1, можно реализовать прецизионное сопровождение траектории объекта ($m_{x'} \rightarrow 0$; $\sigma_{x'} \rightarrow 0$) за счет минимизации аппаратных ошибок измерения координат в канале слежения.

На основании проведенного эксперимента, по исследованию влияния ошибок квантования в следящем контуре цифрового устройства на точность определения координат объекта, можно сделать следующие выводы: введение дополнительного контура слежения по дробной части результата в цифровом измерительном устройстве позволяет существенно (в 5...10 раз) повысить точность измерения координат объекта; для уменьшения аппаратных оши-

бок измерения координат следует обеспечивать значение крутизны дискриминатора в следящем контуре близкое к единице; округление выходного результата траекторного фильтра и управление по этому сигналу генератором стробов более предпочтительно, чем усечение результата, что объясняется отсутствием смещения строба относительно энергетического центра объекта при округлении.

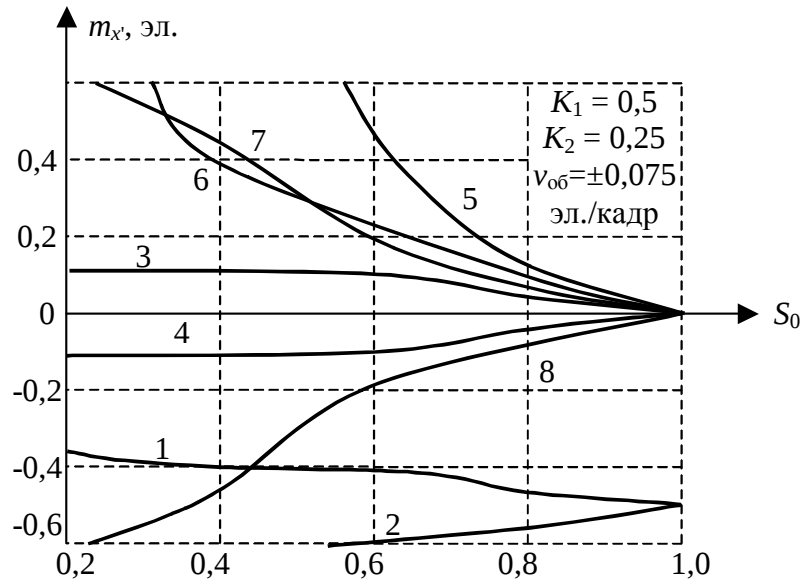


Рис. 1.59. – Зависимости математического ожидания ошибки оценки координат подвижного объекта от крутизны дискриминатора.

Без дробной части: 1 – усечение, $v_{об} \rightarrow$; 2 – усечение, $v_{об} \leftarrow$; 3 – округление, $v_{об} \rightarrow$; 4 – округление, $v_{об} \leftarrow$. С учетом дробной части: 5 – усечение, $v_{об} \rightarrow$; 6 – усечение, $v_{об} \leftarrow$; 7 – округление, $v_{об} \rightarrow$; 8 – округление, $v_{об} \leftarrow$.

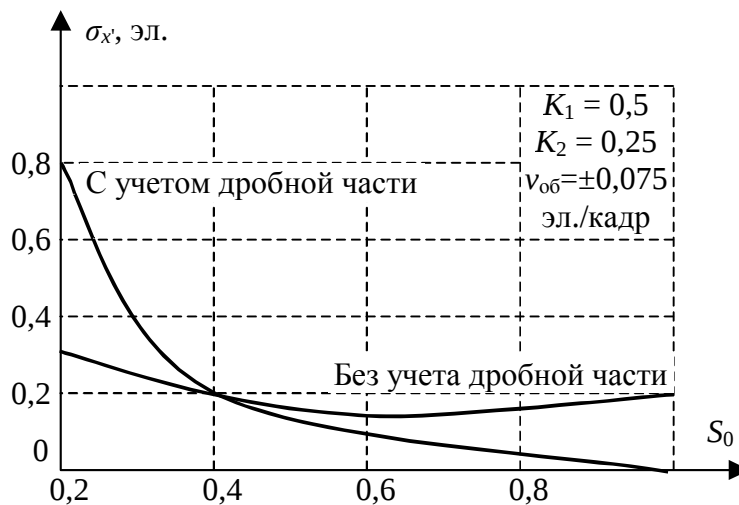


Рис. 1.60. – Зависимости среднего квадратического отклонения ошибки оценки координат подвижного объекта от значений крутизны дискриминатора

2. СИНТЕЗ ЦИФРОВЫХ ФИЛЬТРОВ

2.1. Синтез цифровых устройств для обработки одномерных данных

Процедура синтеза цифровых фильтров (ЦФ) состоит из двух основных этапов: аппроксимации заданных характеристик ЦФ и последующей реализации ЦФ в программном или аппаратном виде. Первый этап предполагает нахождение передаточной функции $H(z)$, удовлетворяющей заданным требованиям. Обычно эти требования выдвигаются прежде всего для частотных характеристик. Передаточная функция $H(z)$ рассчитывается двумя способами: прямым и по аналоговому фильтру-прототипу. В первом случае она определяется непосредственно по заданным параметрам ЦФ, во втором случае вначале находится передаточная функция аналогового фильтра-прототипа (АФП) в p -плоскости, а затем – его дискретная модель в z -плоскости. Методы синтеза по фильтру-прототипу позволяют эффективно использовать богатый опыт проектирования аналоговых фильтров, и поэтому они широко применяются при аппроксимации заданных характеристик ЦФ.

Проектирование ЦФ включает пять основных этапов:

1. Решение аппроксимационной задачи с целью определения коэффициентов a_i , b_j передаточной функции $H(z)$, при которых фильтр удовлетворяет заданным временным либо частотным характеристикам.

2. Выбор структуры или формы реализации цифрового фильтра (прямая, каноническая, каскадная, параллельная, лестничная и т.д.).

3. Расчет разрядности входного и выходного сигналов, разрядности коэффициентов фильтра и арифметических устройств. При этом исключается возможность переполнения разрядности путем введения масштабирующих коэффициентов [1].

4. Проверка моделированием соответствия характеристик разработанного ЦФ заданным. Проверка производится в нереальном времени с «реальным» алгоритмом на ЭВМ. При несоответствии характеристик возвращается сначала к пункту 2, а затем к пункту 1.

5. Построение функциональных и принципиальных схем фильтров, монтаж устройства при аппаратной реализации или программирование ЦФ на языке низкого (ассемблер) или высокого уровня (си, паскаль, фортран и т.д.). При программной реализации ЦФ производится разработка специализированного интерфейса общения с периферийными устройствами ввода-вывода данных. Можно пользоваться стандартными интерфейсами, но, как правило, они имеют ограниченное быстродействие при большой универсальности.

Подобно расчету аналоговых фильтров расчет цифровых фильтров включает в себя процесс нахождения подходящей передаточной функции, которая должным образом удовлетворяет предъявленным требованиям.

В основе синтеза ЦФ лежит преобразование соответствующего аналогового фильтра-прототипа. Такая процедура называется дискретизацией фильтра-прототипа. Таким образом, задачу дискретизации аналоговой передаточной функции $K(p)$ в соответствующую ей цифровую функцию $H(z)$ можно назвать одной из основных в синтезе ЦФ. В этом случае синтез цифровых фильтров состоит из следующих этапов:

1. Получение подходящей передаточной функции аналогового фильтра-прототипа $K(p)$, $h_a(t)$.
2. Создание процедуры перехода, которая преобразует функцию $K(p)$ в соответствующую ей передаточную функцию $H(z)$.

2.2. Синтез ЦФ по методу инвариантного преобразования импульсной характеристики

В основу данного метода положен принцип соответствия ИХ ЦФ выборкам из ИХ аналогового фильтра-прототипа, либо абстрактно заданной ИХ аналоговой цепи – $h_a(nT)$.

Пусть $h_a(t)$ – ИХ АФП, тогда $h(nT) = T \cdot h_a(t)|_{t=nT}$, $n = 0, 1, 2, \dots$.

При $n < 0$, $h(nT) \equiv 0$. Таким образом $\{h(nT)\} = \{h(0), h(T), h(2T), \dots\}$.

Результатом синтеза ИХ данным методом может быть как нерекурсивный, так и рекурсивный цифровые фильтры. Необходимо правильно выбрать интервал дискретизации T . Он определяет степень приближения частотных и временных свойств ЦФ к аналоговому фильтру. Амплитудно-частотная характеристика цифрового фильтра $A(\omega)$ связана с соответствующей характеристикой аналогового фильтра-прототипа $K(\omega)$ соотношением:

$$A(\omega) = \frac{K(0)}{2} + \sum_{k=-\infty}^{\infty} K(\omega - k\omega_d),$$

где ω_d – угловая частота дискретизации. Так как АЧХ АФП на практике не имеет резких ограничений, то в процессе её дискретизации возникает явление наложения (между последовательными членами приведённого выше выражения), представленное графически на рис. 2.1. Пунктиром на рисунке показана суммарная АЧХ, то есть АЧХ ЦФ, которая, как видно из рис. 2.1, не идентична исходной АЧХ аналогового фильтра-прототипа из-за эффекта наложения.

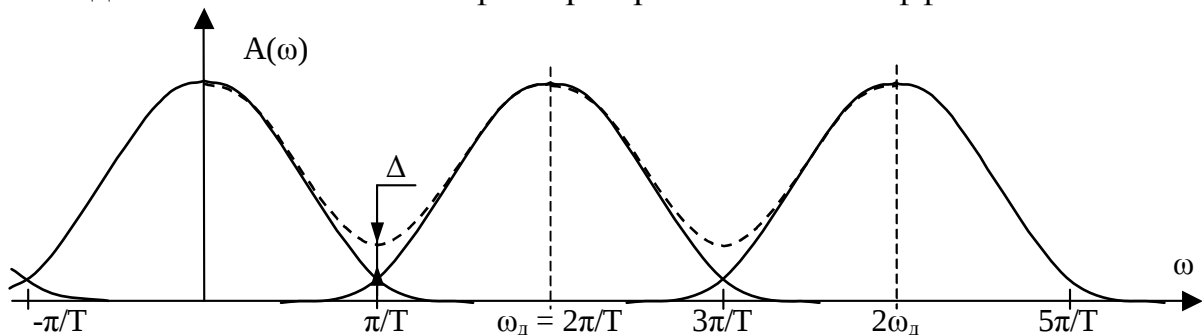


Рис. 2.1. – АЧХ цифрового фильтра

Частота дискретизации ω_d находится из допустимой ошибки перекрытия Δ амплитудно-частотных характеристик фильтра-прототипа: $\omega_d = 2\pi/T$.

Для синтеза нерекурсивного фильтра задаемся ошибкой ε . Число выборок будет определять порядок фильтра M , т.е.

$$\{h(nT)\} = \{Th_a(0), Th_a(T), Th_a(2T), \dots, Th_a(MT - T)\} = \{a_0, a_1, a_2, \dots, a_{M-1}\},$$

$$h(nT) \equiv 0, \quad n < 0.$$

Пусть $h_a(t) = \frac{1}{\tau} e^{-\frac{t}{\tau}}, t \geq 0$, тогда $h(nT) = \frac{T}{\tau} e^{-\frac{nT}{\tau}}, n \geq 0$.

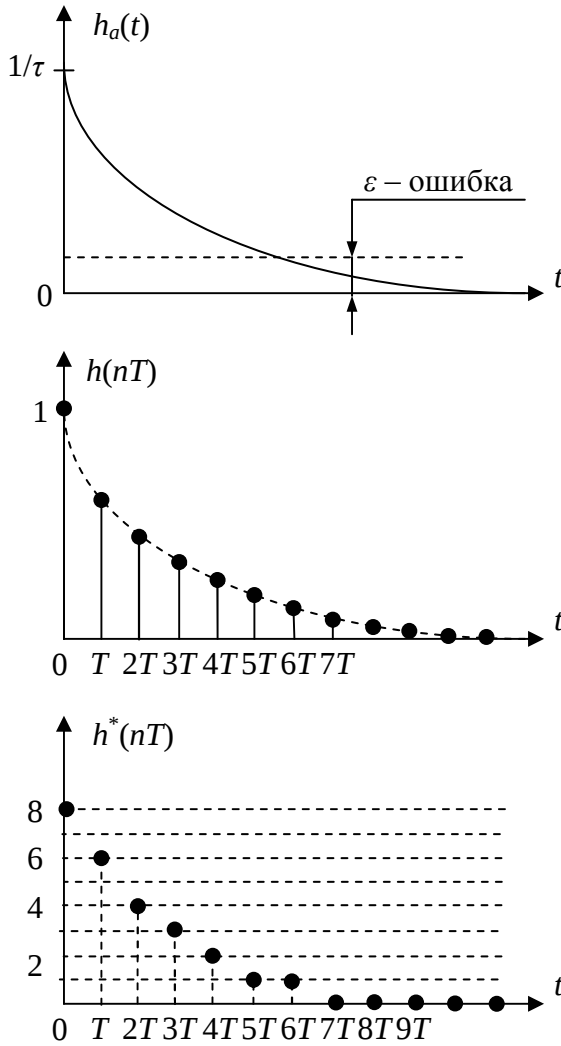


Рис. 2.2. – Дискретизация и квантование ИХ АФП ($K_y = 8$)

Системная функция в этом случае имеет вид:

$$H(z) = 1 + e^{-\frac{T}{\tau}} z^{-1} + e^{-\frac{2T}{\tau}} z^{-2}, \text{ а комплексная частотная характеристика:}$$

$$H(e^{j\omega T}) = 1 + e^{-\frac{T}{\tau}} e^{-j\omega T} + e^{-\frac{2T}{\tau}} e^{-2j\omega T}.$$

Чем меньше интервал дискретизации T , тем лучше степень приближения к ИХ аналогового фильтра-прототипа, но при этом увеличивается порядок фильтра M . Чем больше отношение τ/T , тем больше требуется значимых отсчетов в ИХ, т.е. порядок M нерекурсивного фильтра растет, и тем более нерациональным становится использование нерекурсивного варианта построения ЦФ.

Количество уровней квантования (K_y) определяет значения коэффициентов фильтра и может снижать порядок нерекурсивного фильтра M (при малом K_y).

Преобразование аналоговой ИХ в дискретную и цифровую проиллюстрировано на рис. 2.2.

Ограничимся тремя отсчетами ИХ $h(nT)$ для НРФ и опустим множитель T/τ . Тогда разностное уравнение будет выглядеть следующим образом:

$$y(nT) = e^{-\frac{0}{\tau}} x(nT) + e^{-\frac{T}{\tau}} x(nT - T) + e^{-\frac{2T}{\tau}} x(nT - 2T), \quad n \geq 0.$$

В случае *рекурсивного фильтра* разностное уравнение примет следующий вид: $y(nT) = x(nT) + K \cdot y(nT - T)$, $K = e^{-\frac{T}{\tau}}$, $n \geq 0$.

Системная функция такого фильтра: $H(z) = \frac{z}{z - e^{-\frac{T}{\tau}}}$.

Структурная схема рекурсивного варианта приведена на рис. 2.3.

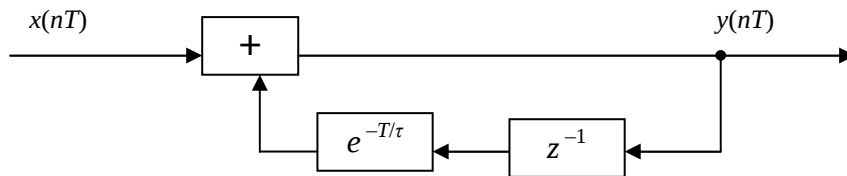


Рис. 2.3. – Сглаживающее звено ($K = e^{-T/\tau}$)

Импульсная характеристика: $h(nT) = K^n = e^{-\frac{nT}{\tau}}$, $n = 0, 1, 2, \dots$.

Комплексная ЧХ: $H(e^{-j\omega T}) = \frac{1}{1 - e^{-\frac{T}{\tau}} e^{-j\omega T}} = \frac{1}{1 - K \cos(\omega T) + jK \sin(\omega T)}$.

АЧХ: $A(\omega) = \frac{1}{\sqrt{1 - 2K \cos(\omega T) + K^2}}$.

ФЧХ: $\varphi(\omega) = \arctg\left(\frac{\sin(\omega T)}{\cos(\omega T) - K^{-1}}\right)$.

Как видно, в отличие от аналогового фильтра-прототипа АЧХ данного ЦФ периодична. Таким образом, соответствие АЧХ аналогового и цифрового фильтров происходит только на главном интервале частот $\Theta_\omega = [0; \pi/T]$ (см. рис. 2.4). Далее частотные характеристики периодически повторяются. ФЧХ ЦФ приведены на рис. 2.5.

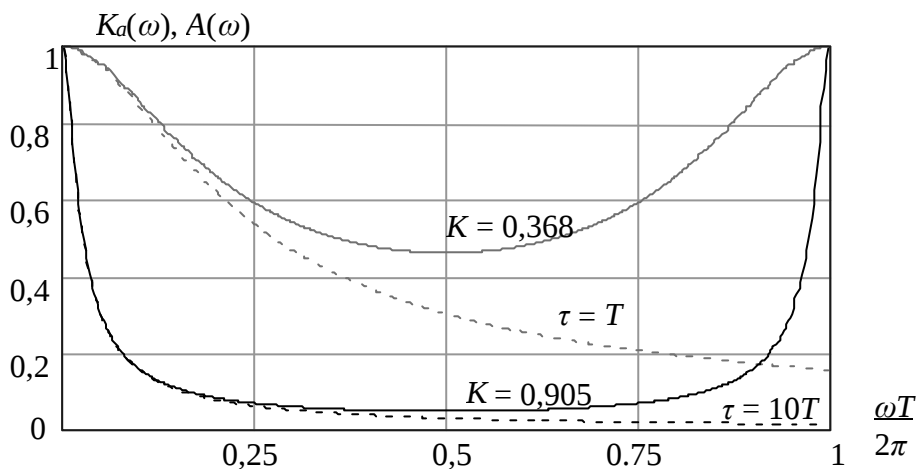


Рис. 2.4. – АЧХ АФП и ЦФ ($\tau = 10T$, $K = 0,905$ и $\tau = T$, $K = 0,368$)

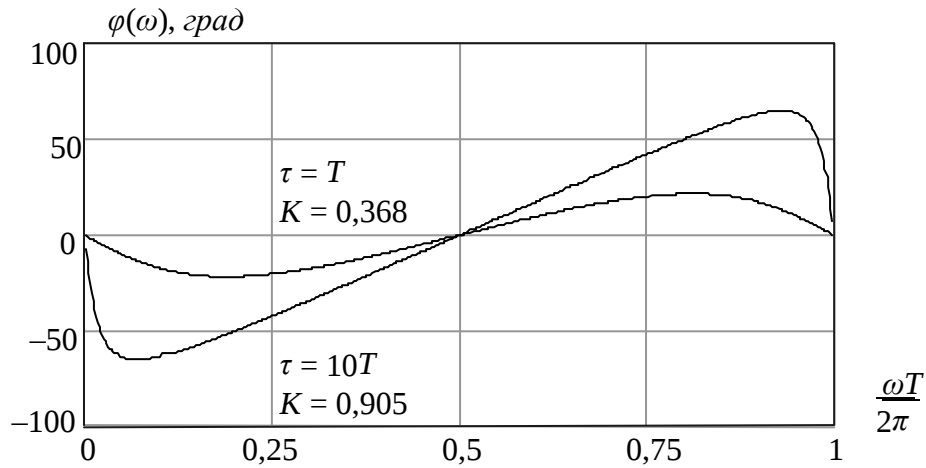


Рис. 2.5. – ФЧХ ЦФ ($\tau = 10T$, $K = 0,905$ и $\tau = T$, $K = 0,368$)

Вычислим нормированные АЧХ для рекурсивного и нерекурсивного вариантов фильтра при $\tau = 5T$.

Аналоговый фильтр-прототип (АФП):

$$\left| \frac{K(j\omega)}{K(0)} \right| = \frac{1}{\sqrt{1 + \omega^2 \tau^2}} = \frac{1}{\sqrt{1 + 25\omega^2 T^2}};$$

Цифровой нерекурсивный фильтр (ЦНРФ):

$$\left| \frac{H_1(e^{j\omega T})}{H(1)} \right| = \frac{\sqrt{2,12 + 2,735 \cos(\omega T) + 1,34 \cos(2\omega T)}}{2,489}.$$

Цифровой рекурсивный фильтр (ЦРФ):

$$\left| \frac{H_2(e^{j\omega T})}{H(1)} \right| = \frac{0,1811}{\sqrt{1,6703 - 1,6375 \cos(\omega T)}};$$

Сведем результаты в табл. 2.1.

Таблица 2.1

Нормированные АЧХ для различных вариантов фильтров при $\tau = 5T$

ωT	АФП	ЦНРФ	ЦРФ
0	1,0000	1,0000	1,0000
0,5	0,3714	0,9201	0,3754
1,0	0,1961	0,7005	0,2046
1,5	0,1322	0,3991	0,1454
2,0	0,0995	0,1308	0,1182
2,5	0,0797	0,2233	0,1050
3,0	0,0665	0,3359	0,1000

Из таблицы видно, что при трех отсчетах ИХ амплитудно-частотные характеристики АФП и ЦНРФ сильно различаются, в то же время, совпадение АЧХ для ЦРФ хорошее. Подъем АЧХ ЦНРФ вблизи точки $\omega T = \pi$ идет из-за члена $\cos(2\omega T)$. Вид АЧХ показан на рис. 2.6.

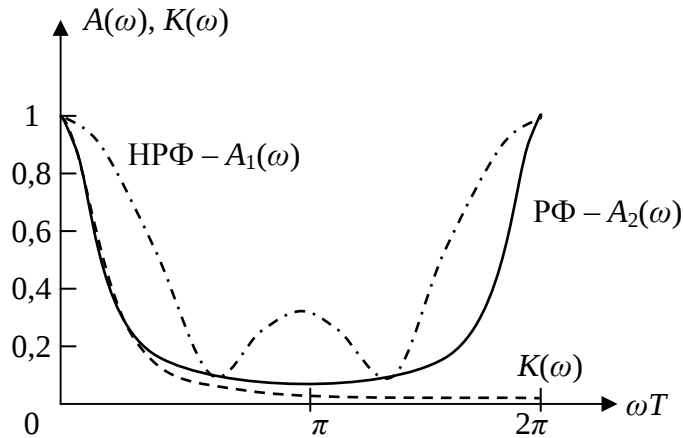


Рис. 2.6. – АЧХ рекурсивного и нерекурсивного вариантов ЦФ ($\tau = 5T$)

Хорошее совпадение по АЧХ наблюдается, если учитывается не менее 20 отсчетов ИХ, а не трех.

Пример 1. Синтез цифрового фильтра второго порядка, эквивалентного колебательному контуру, т.е. синтез цифрового резонатора.

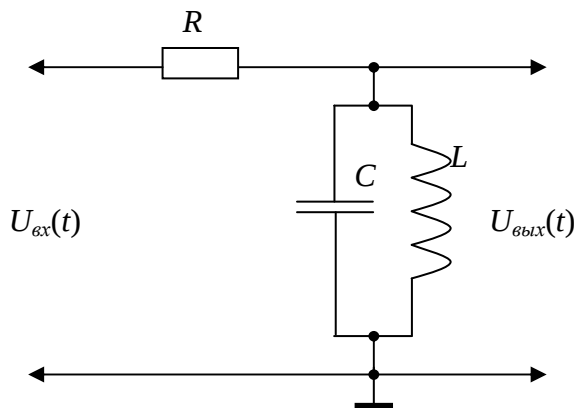


Схема резонансной цепи приведена на рис. 2.7.

$$K(p) = \frac{2\alpha\rho}{p^2 + 2\alpha\rho + \omega_0^2}, \text{ где } \omega_0 = \frac{1}{\sqrt{LC}};$$

$$\rho = \sqrt{\frac{L}{C}}; \quad Q = \frac{R}{\rho}; \quad \alpha = \frac{\omega_0}{2Q} = \frac{1}{2RC};$$

$$\omega_c = \sqrt{\omega_0^2 - \alpha^2} = \omega_0 \sqrt{1 - \frac{1}{4Q^2}};$$

Рис. 2.7. – Резонансный фильтр

α – коэффициент затухания;

ω_0 – резонансная частота;

Q – добротность;

ρ – характеристическое сопротивление;

ω_c – частота свободных колебаний;

$\xi = \frac{\alpha}{\omega_0} = \frac{1}{2Q}$ – коэффициент демпфирования;

$Q = \frac{1}{2}$, ($\xi = 1$) – критический режим;

$Q > \frac{1}{2}$, ($\xi < 1$) – колебательный режим;

$Q < \frac{1}{2}$, ($\xi > 1$) – апериодический режим.

Импульсная характеристика резонансного фильтра в колебательном режиме:

$$h_a(t) = L^{-1}\{K_a(p)\} = 2\alpha e^{-\alpha t} \left[\cos(\omega_c t) - \frac{\alpha}{\omega_c} \sin(\omega_c t) \right], \quad t \geq 0.$$

1. Получаем импульсную характеристику ЦФ:

$$h(nT) = Th_a(t)|_{t=nT} = 2\alpha T e^{-\alpha nT} \left[\cos(\omega_c nT) - \frac{\alpha}{\omega_c} \sin(\omega_c nT) \right], \quad n \geq 0.$$

2. Используя прямое Z-преобразование, получим выражение для системной функции $H(z)$:

$$H(z) = Z\{h(nT)\} = \sum_{n=0}^{\infty} h(nT) z^{-n}.$$

$$\text{Представим } \cos(\omega_c nT) = \frac{e^{j\omega_c nT} + e^{-j\omega_c nT}}{2}, \quad \sin(\omega_c nT) = \frac{e^{j\omega_c nT} - e^{-j\omega_c nT}}{2j} \quad \text{и,}$$

используя формулу для суммы членов геометрической прогрессии

$$\sum_{n=0}^{\infty} q^n = \frac{1}{1-q}, \quad |q| < 1, \quad q_{1,2} = e^{-\alpha T \pm j\omega_c T} \cdot z^{-1}, \quad \text{получим выражение для СФ:}$$

$$H(z) = k_0 \frac{a_0 + a_1 z^{-1}}{1 - b_1 z^{-1} - b_2 z^{-2}}, \quad \text{где } k_0 = 2\alpha T;$$

$$a_0 = 1; \quad a_1 = -e^{-\alpha T} \left(\cos(\omega_c T) + \frac{\alpha}{\omega_c} \sin(\omega_c T) \right);$$

$$b_1 = 2e^{-\alpha T} \cos(\omega_c T); \quad b_2 = -e^{-2\alpha T}.$$

Разностное уравнение данного цифрового резонатора запишем в виде:

$$y(nT) = k_0 [a_0 x(nT) + a_1 x(nT - T)] + b_1 y(nT - T) + b_2 y(nT - 2T), \quad n \geq 0.$$

Структурная схема цифрового резонатора приведена на рис. 2.8.

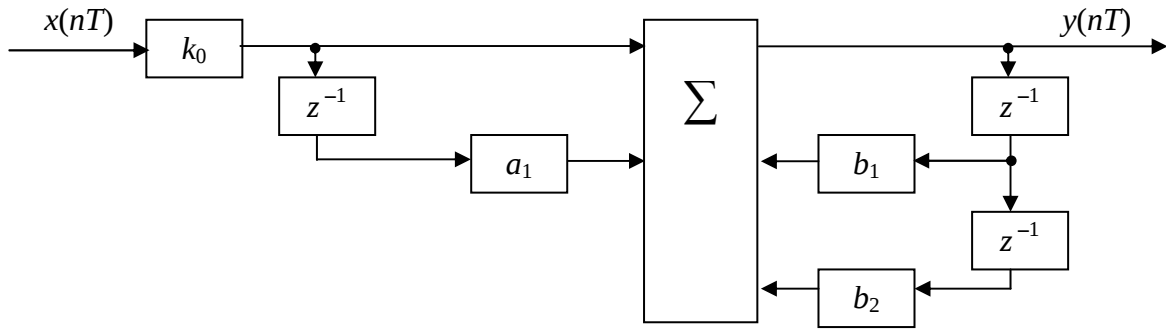


Рис. 2.8. – Структурная схема цифрового резонатора ($Q > 10$)

При $Q > 10$ и $\frac{\alpha}{\omega_c} < 0,05$ коэффициент $a_1 \approx -e^{-\alpha T} \cos(\omega_c T)$.

Рассмотрим расположение особых точек системной функции $H(z)$ в z -плоскости. Системная функция имеет один нуль z_0 и два комплексно-сопряженных полюса z_1 и z_2 (рис. 2.9).

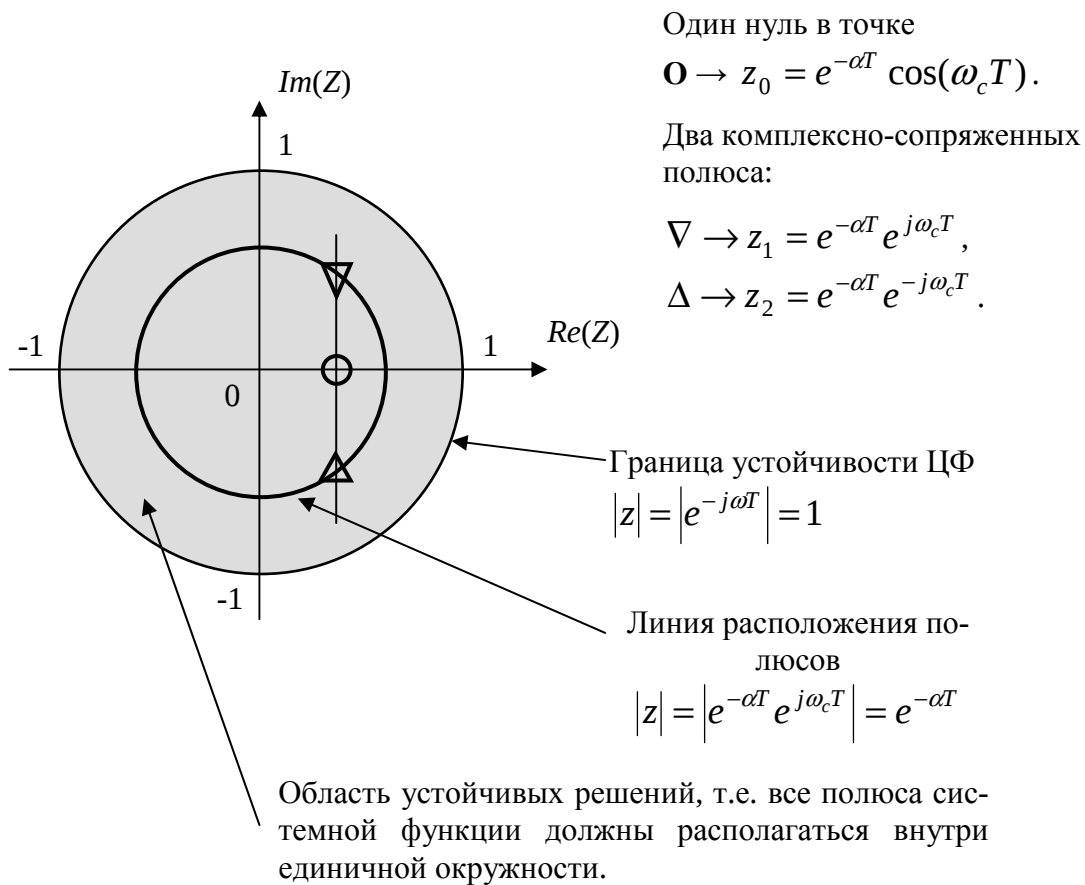


Рис. 2.9. – Расположение особых точек системной функции ЦФ на z -плоскости

При $\alpha \rightarrow 0$ z_1 и z_2 стремятся к границе устойчивости, а при $\alpha \rightarrow \infty$ z_1 и z_2 стремятся к нулю. В последнем случае в системе будут наблюдаться сильно затухающие колебания при ее возбуждении единичным импульсом.

При $\omega_c \rightarrow 0$ z_0, z_1, z_2 стремятся к точке $e^{-\alpha T}$, лежащей на линии $Re(z)$.

При $\omega_c \geq \frac{\pi}{T} = \frac{\omega_0}{2}$ – полюса меняются друг с другом.

Пример 2. Синтез цифрового режекторного фильтра.

Операторный коэффициент передачи аналогового режекторного фильтра:

$$K_{\text{рез}}(p) = \frac{p^2 + \omega_0^2}{p^2 + 2\alpha p + \omega_0^2}, \text{ где } \omega_0 = \frac{1}{\sqrt{LC}}; \rho = \sqrt{\frac{L}{C}}; \alpha = \frac{\omega_0}{2Q};$$

$$\omega_c = \sqrt{\omega_0^2 - \alpha^2} = \omega_0 \sqrt{1 - \frac{1}{4Q^2}}; Q = \frac{\omega_0}{\Delta\omega}; \Delta\omega - \text{полоса частот на уровне } -3 \text{ дБ}$$

или 0,707.

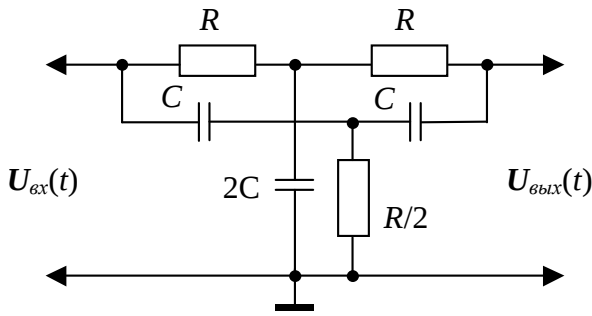


Рис. 2.10. – Аналоговый режекторный фильтр

На рис. 2.10 приведена схема варианта построения аналогового режекторного фильтра.

$$\omega_0 = \frac{1}{RC}; \alpha = \frac{2}{RC};$$

$$Q = \frac{\omega_0}{2\alpha} = \frac{1}{4} < \frac{1}{2} - \text{апериодический режим.}$$

Операторный коэффициент пе-

редачи двойного Т-моста имеет вид:

$$K_{\text{рез}}(p) = \frac{p^2 + \omega_0^2}{p^2 + 2\alpha p + \omega_0^2} \rightarrow K_{\text{рез}}(p) = \frac{p^2 + \frac{1}{(RC)^2}}{p^2 + \frac{4}{RC}p + \frac{1}{(RC)^2}}.$$

В прямую продискретизировать режекторный фильтр не удастся. Поэтому воспользуемся соотношением, справедливым как для аналоговых, так и для цифровых фильтров:

$$K_{\text{рез}}(p) + K_{\text{рез}}(p) = 1,$$

$$H_{\text{рез}}(z) + H_{\text{рез}}(z) = 1.$$

Из последнего выражения следует $H_{\text{рез}}(z) = 1 - H_{\text{рез}}(z) = \frac{a_0' + a_1'z^{-1} + a_2'z^{-2}}{1 - b_1'z^{-1} - b_2'z^{-2}},$

где $a_0' = 1 - k_0 a_0 = 1 - 2\alpha T;$

$$a_1' = -k_0 a_1 - b_1 = -2(1 - \alpha T) e^{-\alpha T} \left(\cos(\omega_c T) + \frac{\alpha}{\omega_c} \sin(\omega_c T) \right);$$

$$a_2' = -b_2 = e^{-2\alpha T};$$

$$b_1' = b_1 = 2e^{-\alpha T} \cos(\omega_c T);$$

$$b_2' = b_2 = -e^{-2\alpha T}.$$

Как видно, данный фильтр имеет на одну задержку в прямой ветви больше, т.к. $a_2' \neq 0$.

Задавая добротность Q , частоту резонанса ω_0 и частоту дискретизации $\omega_d = 5 \div 10 \omega_0 \left(T = \frac{2\pi}{\omega_d} \right)$, вычисляем коэффициент демпфирования ξ , затухание $\alpha = \omega_0 \xi$ и частоту свободных колебаний ω_c . Вычислив указанные параметры, найдем коэффициенты k_0, a_0, a_1, b_1, b_2 для цифрового резонатора и $a_0', a_1', a_2', b_1', b_2'$ для цифрового режекторного фильтра.

2.3. Синтез цифровых фильтров методом отображения дифференциалов

В основе метода лежит дискретизация аналогового фильтра-прототипа, заключающаяся в замене дифференциалов в его дифференциальном уравнении прямыми, обратными или центрированными разностями. Дифференциальное уравнение для аналогового фильтра можно получить непосредственно либо из операторного коэффициента передачи [5].

Любую физически реализуемую цепь можно описать операторным коэффициентом передачи в виде отношения полиномов

$$K(p) = \frac{C(p)}{Q(p)} = \frac{c_0 + c_1 p + c_2 p^2 + \dots + c_M p^M}{q_0 + q_1 p + q_2 p^2 + \dots + q_N p^N}, \text{ где } M \leq N.$$

Из определения операторного коэффициента передачи $K(p) = \frac{Y(p)}{X(p)}$ следует: $Q(p)Y(p) = C(p)X(p)$. Заменяв оператор p на оператор дифференцирования, а изображения $Y(p)$ и $X(p)$ на сигналы $y(t)$ и $x(t)$, получим дифференциальное уравнение, описывающее аналоговый фильтр-прототип:

$$\sum_{j=0}^N q_j \frac{y^j(t)}{dt^j} = \sum_{i=0}^M c_i \frac{x^i(t)}{dt^i}, \quad t \geq 0.$$

Простейший метод дискретизации системы состоит в замене дифференциалов на прямые или обратные разности (чаще используются обратные разности):

$\nabla_0[f(nT)] = f(nT)$ – нулевая разность;

$\nabla_1[f(nT)] = \frac{1}{T}[f(nT) - f(nT - T)]$ – первая разность;

$\nabla_2[f(nT)] = \frac{1}{T}[\nabla_1[f(nT)] - \nabla_1[f(nT - T)]] =$
 $= \frac{1}{T^2}[f(nT) - 2f(nT - T) + f(nT - 2T)]$ – вторая разность и т. д.

Структурная схема цифрового фильтра для формирования второй разности приведена на рис. 2.11.

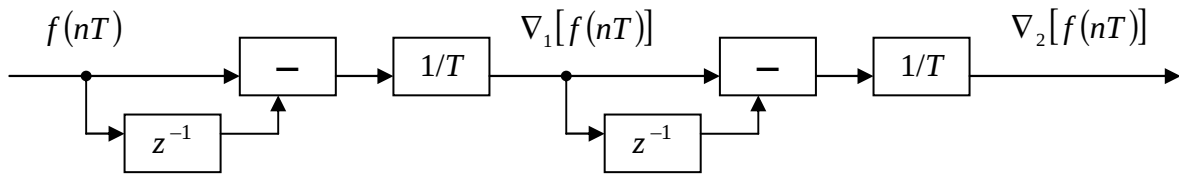


Рис. 2.11. – Формирователь второй разности

Заменяя в дифференциальном уравнении, описывающем АФП, дифференциалы соответствующими обратными разностями, получим неприведенное разностное уравнение:

$$\sum_{j=0}^N q_j \nabla_j[y(nT)] = \sum_{i=0}^M c_i \nabla_i[x(nT)], \quad n \geq 0.$$

Затем приводим РУ к стандартному виду:

$$y(nT) = \sum_{i=0}^M a_i x(nT - iT) + \sum_{j=1}^N b_j y(nT - jT), \quad n \geq 0.$$

Таким образом, при использовании обратных разностей производится следующая замена:

$$\frac{df(t)}{dt} \leftrightarrow \frac{f(nT) - f(nT - T)}{T},$$

что с точки зрения операторов соответствует соотношениям

$$p = \frac{1 - z^{-1}}{T} \quad \text{и} \quad z = \frac{1}{1 - pT}.$$

Используя указанные соотношения можно непосредственно перейти от операторного коэффициента к системной функции и обратно:

$$H(z) = K(p) \Big|_{p = \frac{1 - z^{-1}}{T}}; \quad K(p) = H(z) \Big|_{z = \frac{1}{1 - pT}}.$$

Полученную системную функцию приводим к стандартному виду:

$$H(z) = \frac{\sum_{i=0}^M a_i z^{-i}}{1 - \sum_{j=1}^N b_j z^{-j}}.$$

Пример 1: Цифровой режекторный фильтр.

Синтезируем цифровой режекторный фильтр по заданному аналоговому фильтру-прототипу (рис. 2.12).

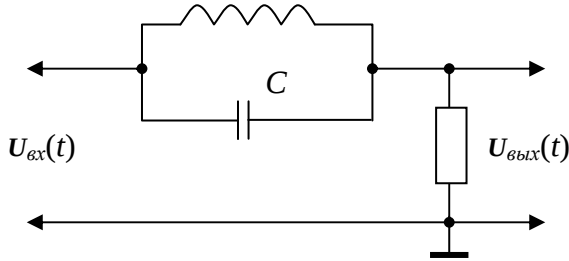


Рис. 2.12. – Режекторный фильтр-прототип

$$K_{\text{реж}}(p) = \frac{p^2 + \omega_0^2}{p^2 + 2\alpha p + \omega_0^2}, \text{ где}$$

$$\omega_0 = \frac{1}{\sqrt{LC}}, \alpha = \frac{\omega_0}{2Q} = \frac{1}{2RC}.$$

Заменяя оператор p на оператор $\frac{d}{dt}$, а $X(p) \rightarrow x(t)$ и $Y(p) \rightarrow y(t)$, получим дифференциальное уравнение:

$$(p^2 + 2\alpha p + \omega_0^2)Y(p) = (p^2 + \omega_0^2)X(p),$$

$$\frac{d^2 y(t)}{dt^2} + 2\alpha \frac{dy(t)}{dt} + \omega_0^2 y(t) = \frac{d^2 x(t)}{dt^2} + \omega_0^2 x(t), \quad t \geq 0.$$

Запишем разностное уравнение в обратных разностях:

$$\nabla_2[y(nT)] + 2\alpha \nabla_1[y(nT)] + \omega_0^2 y(nT) = \nabla_2[x(nT)] + \omega_0^2 x(nT), \quad n \geq 0.$$

Раскроем выражение и запишем разностное уравнение:

$$(1 + 2\alpha T + \omega_0^2 T^2)y(nT) - 2(1 + \alpha T)y(nT - T) + y(nT - 2T) = \\ = (1 + \omega_0^2 T^2)x(nT) - 2x(nT - T) + x(nT - 2T), \quad n \geq 0.$$

Выражая это уравнение относительно $y(nT)$, найдем коэффициенты ЦФ:

$$a_0 = \frac{1 + \omega_0^2 T^2}{1 + 2\alpha T + \omega_0^2 T^2}; \quad a_1 = -\frac{2}{1 + 2\alpha T + \omega_0^2 T^2}; \quad a_2 = \frac{1}{1 + 2\alpha T + \omega_0^2 T^2};$$

$$b_1 = \frac{2(1 + \alpha T)}{1 + 2\alpha T + \omega_0^2 T^2}; \quad b_2 = -\frac{1}{1 + 2\alpha T + \omega_0^2 T^2}.$$

Достоинство метода замены дифференциалов разностями заключается в простоте перехода от операторного коэффициента передачи $K(p)$ непосредственно к системной функции $H(z)$.

Недостаток метода заключается в плохом сохранении частотных характеристик аналогового фильтра-прототипа независимо от использования прямых или обратных разностей. Необходимо, как и в методе инвариантного преобразования ИХ, выбрать интервал дискретизации T достаточно малым, чтобы передать все особенности ИХ, или правильно продискретизировать дифференциальное уравнение.

Рассмотрим, как отображается левая p -полуплоскость в z -плоскость:

замена $p = \frac{1-z^{-1}}{T}$, $z = \frac{1}{1-pT}$; частотная ось $p = j\omega_a$,

$$z = \frac{1}{1-j\omega_a T} = \frac{1}{2} \left(1 + \frac{1+j\omega_a}{1-j\omega_a} \right) = \frac{1}{2} (1 + e^{2j \arctg(\omega_a T)}),$$

$$\operatorname{Re}[z] = \frac{1}{2} [1 + \cos(2 \arctg(\omega_a T))], \quad \operatorname{Im}[z] = \frac{1}{2} \sin(2 \arctg(\omega_a T)),$$

$$\left\{ \operatorname{Re}[z] - \frac{1}{2} \right\}^2 + \{\operatorname{Im}[z]\}^2 = \left(\frac{1}{2} \right)^2 - \text{уравнение окружности в } z\text{-плоскости.}$$

На рис. 2.13. показано, что совпадение аналоговых ω_a и ω_u частот происходит на небольшом участке $\omega T \ll \pi$.

При равной частоте дискретизации метод отображения дифференциалов работает хуже, чем метод инвариантного преобразования ИХ. С увеличением частоты дискретизации совпадение характеристик улучшается, но реализация ЦФ становится затруднительной, т.к. требуется повышенное быстродействие ЦФ.

Метод отображения дифференциалов можно улучшить, если использовать разности более высоких порядков для замены дифференциалов низких порядков или использовать центрированные разности. Однако из-за сложности замены такой подход не нашел практического применения.

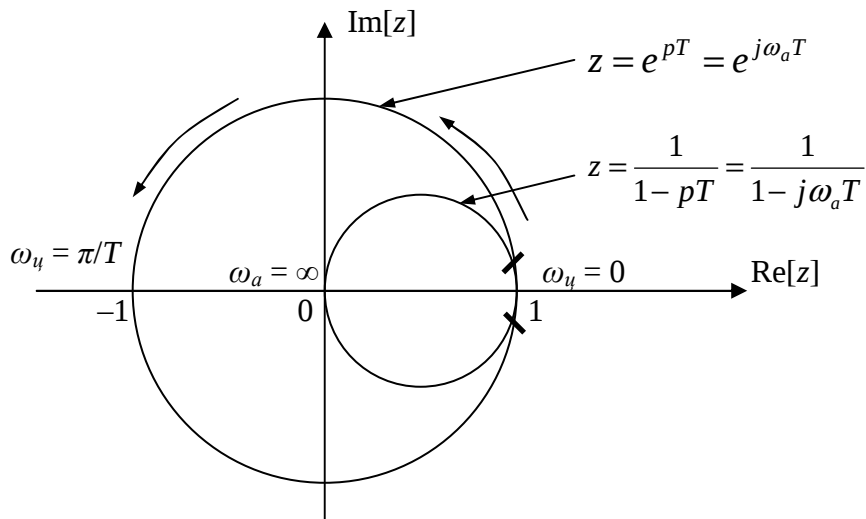


Рис. 2.13. – Отображение оси $j\omega_a$ из p -плоскости в z -плоскость

2.4. Синтез цифровых фильтров методом билинейного преобразования

Синтез цифрового фильтра производится по операторному коэффициенту передачи аналогового фильтра-прототипа. Метод несколько сложнее предыдущего, но дает лучшие результаты.

Исходные данные: передаточная функция аналогового фильтра-прототипа в операторной форме $K(p) = \frac{c_0 + c_1 p + c_2 p^2 + \dots + c_M p^M}{d_0 + d_1 p + d_2 p^2 + \dots + d_N p^N} = \frac{C(p)}{D(p)}$ есть дробно-рациональное выражение, описывающее любую физически реализуемую цепь [2].

Частотная характеристика цифрового фильтра периодична (рис. 2.14).

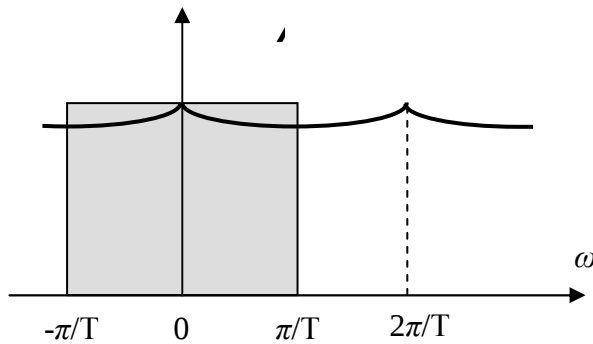


Рис. 2.14 – Периодичность АЧХ

Используем основной интервал частот $\omega \in \left[-\frac{\pi}{T}; \frac{\pi}{T}\right]$. Системную

функцию $H(z)$ можно найти заменой

$$e^{pT} = z \rightarrow pT = \ln(z) \rightarrow p = \frac{\ln(z)}{T}.$$

Точно реализовать частотные свойства аналогового фильтра-прототипа нельзя, т.к.

$$\omega_{a\phi} \in [-\infty; \infty], \quad \omega_{ц\phi} \in \left[-\frac{\pi}{T}; \frac{\pi}{T}\right].$$

При билинейном преобразовании (БЛП) происходит трансформация частот (рис. 2.15).

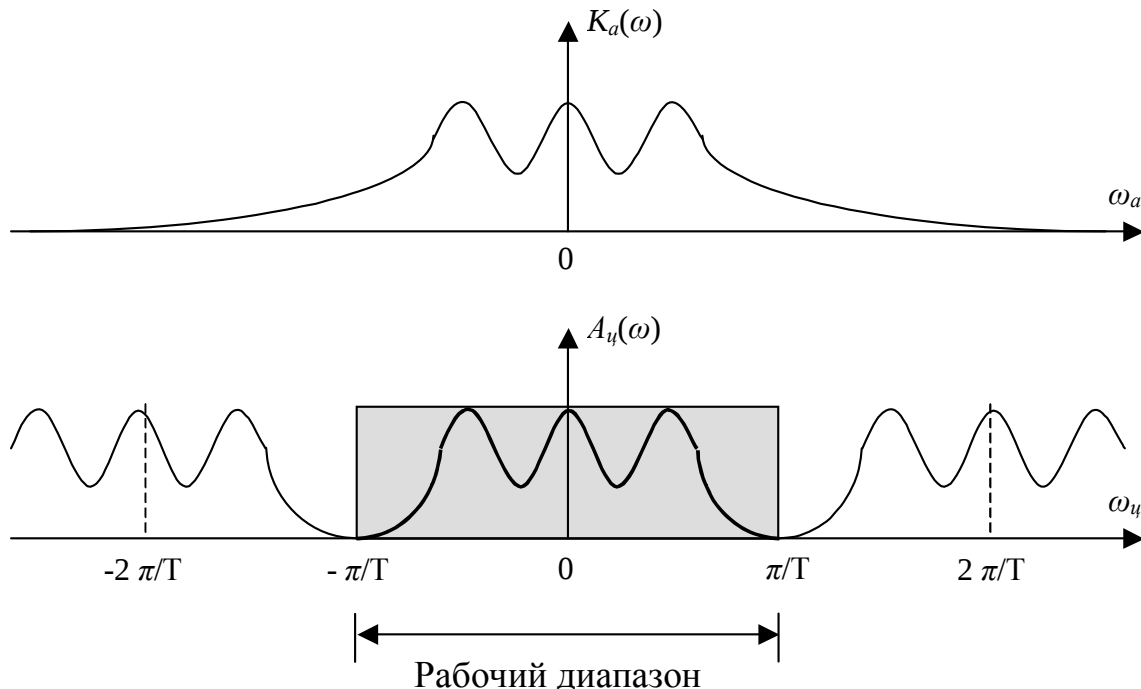


Рис. 2.15. – Трансформация частот при билинейном преобразовании

Произошла трансформация частоты $-\infty$ в $-\pi/T$, а ∞ в π/T . Таким образом, произошло сжатие полосы частот в цифровом фильтре.

Алгоритм использования метода билинейного преобразования:

1. Разложим $p = \frac{1}{T} \ln(z)$ в ряд $p = \frac{2}{T} \left(\zeta + \frac{1}{3} \zeta^3 + \frac{1}{5} \zeta^5 + \dots \right)$, $\zeta = \frac{z-1}{z+1}$.
2. Ограничимся первым членом ряда $p = \frac{2}{T} \zeta = \frac{2}{T} \cdot \frac{z-1}{z+1}$.
3. Производим коррекцию характерной частоты аналогового фильтра-прототипа (частоты резонанса, режекции или среза) $\omega_a = \frac{2}{T} \operatorname{tg} \left(\frac{\omega_y T}{2} \right)$.
4. Делая замену (см. пункт 2), приводим выражение передаточной функ-

ции аналогового фильтра $K(p) = \frac{C(p)}{D(p)}$ к виду $H(z) = \frac{\sum_{i=0}^M a_i z^{-i}}{1 - \sum_{j=1}^N b_j z^{-j}}$.

5. По полученному выражению строим цифровой фильтр. При этом, если аналоговый прототип устойчив ($\operatorname{Re}(p_{\text{полюс}}) < 0$), то и соответствующий ему ЦФ также будет устойчив.

В качестве примера определим структуру цифрового интегратора.

Пусть $K(p) = p^{-1}$. Тогда $p^{-1} = \frac{T}{2} \frac{z+1}{z-1} = \frac{T}{2} \left(\frac{z}{z-1} + \frac{z \cdot z^{-1}}{z-1} \right)$.

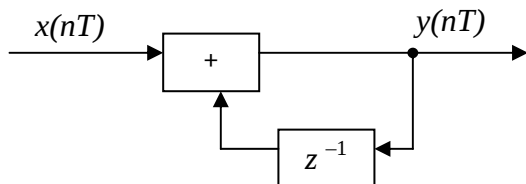


Рис. 2.16. – Накапливающий сумматор

Ранее мы определили операцию суммирования с накоплением в виде (рис. 2.16):

$$y(nT) = x(nT) + y(nT-T); \quad H_u(z) = \frac{z}{z-1},$$

по сути, это цифровой сумматор с накоплением результата (накапливающий сумматор). Тогда цифровой интегратор

будет иметь вид: $H(z) = \left[H_u(z) + H_u(z) \cdot z^{-1} \right] \cdot \frac{T}{2}$.

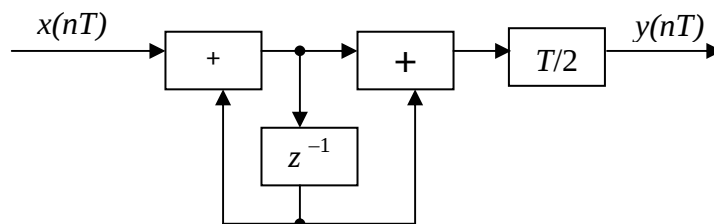


Рис. 2.17. – Цифровой интегратор, синтезированный методом БЛП

Связь частот ω_u и ω_a при БЛП (коррекция характерной частоты)

Связь характерных точек при конформном отображении левой полуплоскости $p = \sigma + j\omega$ в единичный круг в z -плоскости приведена в таблице 2.2 и рис. 2.18 ($T = 1$ с).

Таблица 2.2

Конформное отображение $z = e^{pT}$; $p = \sigma + j\omega$

	σ	ω	z
1.	0	0	1
2.	0	π	-1
3.	0	$\pi/2$	j
4.	0	$-\pi/2$	$-j$
5.	$-\infty$	любое	0
6.	$-\alpha$	$-\pi < \omega \leq \pi$	$ z = e^{-\alpha} \rightarrow$ окружность радиуса $e^{-\alpha}$

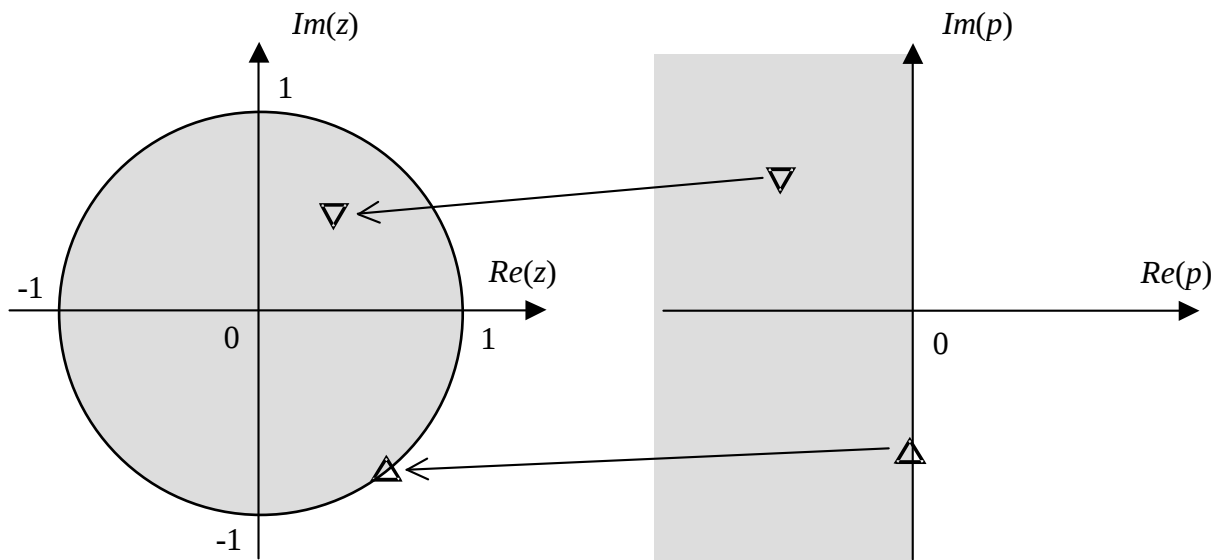


Рис. 2.18. – Левая p -полуплоскость отображается в единичный круг ($Re(p) < 0$)

При билинейном преобразовании левая p -полуплоскость также отображается в единичный круг в z -плоскости. Связь цифровых и аналоговых частот будет нелинейной:

$$p = \frac{2}{T} \frac{z-1}{z+1} \bigg|_{z=e^{j\omega T}} \Rightarrow j\omega_a = \frac{2(e^{j\omega_u T} - 1)}{T(e^{j\omega_u T} + 1)} = j \frac{2}{T} \operatorname{tg} \left(\frac{\omega_u T}{2} \right).$$

$$\text{Отсюда } \omega_a = \frac{2}{T} \operatorname{tg} \left(\frac{\omega_u T}{2} \right).$$

Видно, что при $T \rightarrow 0$ ($f_{\text{дискр}} \rightarrow \infty$), $\omega_a \approx \frac{2}{T} \cdot \frac{\omega_u T}{2} = \omega_u$.

Рассмотрим, как при билинейном преобразовании происходит отображение оси частот $p = j\omega_a$ из p -плоскости в z -плоскость.

Преобразуем формулу:

$$p = \frac{2}{T} \cdot \frac{z-1}{z+1} \rightarrow z = \frac{2+pT}{2-pT} = \frac{2+j\omega_a T}{2-j\omega_a T} = e^{2j \operatorname{arctg}\left(\frac{\omega_a T}{2}\right)}.$$

Выделим действительную и мнимую части z :

$$\operatorname{Re}[z] = \cos\left[2 \operatorname{arctg}\left(\frac{\omega_a T}{2}\right)\right], \quad \operatorname{Im}[z] = \sin\left[2 \operatorname{arctg}\left(\frac{\omega_a T}{2}\right)\right].$$

Видно, что $\{\operatorname{Re}[z]\}^2 + \{\operatorname{Im}[z]\}^2 = 1$, $-\infty < \omega_a < \infty$. Это уравнение единичной окружности в z -плоскости.

Определяя $z = e^{j\omega_y T}$ и используя приведенное выше соотношение, получим формулу трансформации частот при билинейном преобразовании:

$$j\omega_y T = 2j \cdot \operatorname{arctg}\left(\frac{\omega_a T}{2}\right) \Rightarrow \omega_y = \frac{2}{T} \operatorname{arctg}\left(\frac{\omega_a T}{2}\right).$$

Из данной формулы видно, что частота $\omega_a = \infty$ соответствует цифровой частоте $\omega_y = \frac{2}{T} \cdot \frac{\pi}{2} = \frac{\pi}{T}$, а $\omega_a = -\infty$ соответствует $\omega_y = -\frac{\pi}{T}$.

Вывод. Отображение $z = e^{pT}$, как и отображение при БЛП $z = \frac{2+pT}{2-pT}$, преобразует левую p -полуплоскость в единичный круг в z -плоскости, а частотная ось $p = j\omega_a$ преобразуется в единичную окружность, только связь частот ω_y и ω_a при БЛП будет *нелинейная*. Однако лучшего отображения бесконечного интервала частот $\omega_a \in (-\infty, \infty)$ в конечный интервал $\omega_y \in \left(-\frac{\pi}{T}, \frac{\pi}{T}\right)$ для практического синтеза цифровых фильтров пока не найдено.

На рис. 2.19 показана связь аналоговых и цифровых частот, рассчитанная по приведенным соотношениям. Из рисунка видно, что с уменьшением интервала дискретизации T область примерного линейного соответствия цифровых и аналоговых частот увеличивается. Так, при $T = 1$ с, $\omega_a \in [-2, 2]$, а при $T = 0,2$ с $\omega_a \in [-10, 10]$.

На рис. 2.20 показано, как изменяется вид АЧХ аналогового фильтра-прототипа при билинейном преобразовании.

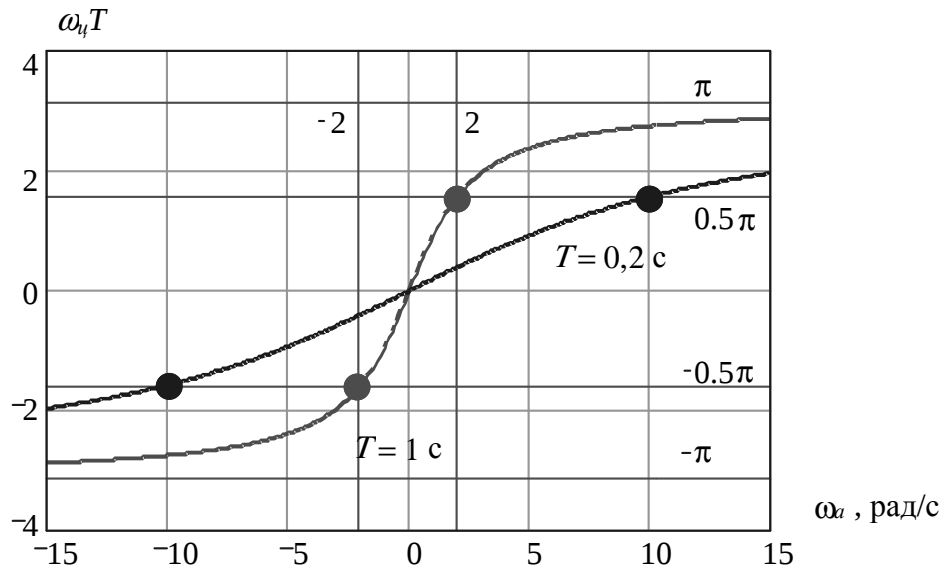


Рис. 2.19. – Связь аналоговых и цифровых частот при БЛП

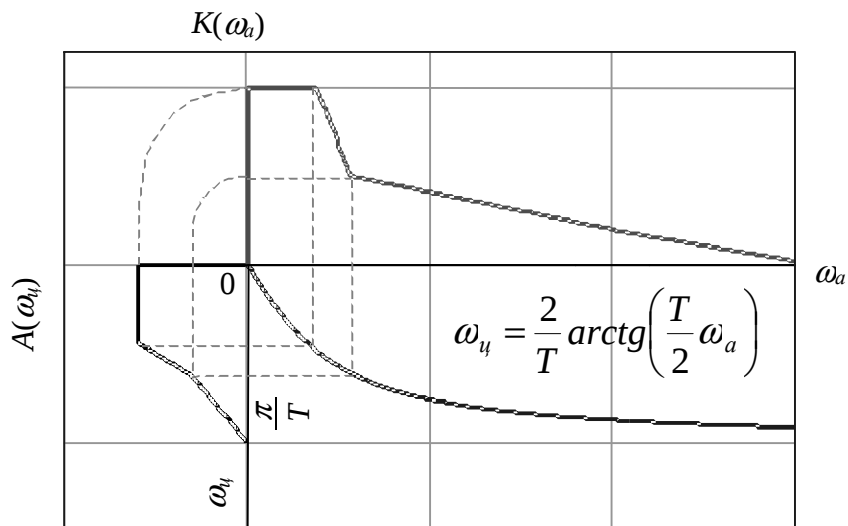


Рис. 2.20. – Трансформация АЧХ при БЛП

Соотношение $\omega_a = \frac{2}{T} \operatorname{tg}\left(\frac{\omega_d T}{2}\right)$ позволяет скорректировать частотный

масштаб при проектировании ЦФ. Пусть, например, требуется ЦФ с заданной частотной характеристикой. Для решения задачи синтеза такого фильтра необходимо так подобрать аналоговый фильтр-прототип, чтобы его частотная характеристика не совпадала точно с ЧХ проектируемого фильтра, а была деформирована по оси частот в соответствии с приведенным выше выражением.

Если известна частотная характеристика $K(\omega)$ или передаточная функция $K(p)$ такого фильтра-прототипа, то с помощью преобразования $p = \frac{2}{T} \frac{z-1}{z+1}$

находят системную функцию $H(z)$ цифрового фильтра и определяют его структурную схему.

Рассмотренный метод билинейного преобразования является одним из основных часто используемых на практике методов синтеза цифровых фильтров [2].

Пример. Требуется спроектировать ЦФ низких частот, имеющий частоту дискретизации $f_{\text{дискр}} = 10$ кГц. АЧХ ЦФ должна быть максимально плоской в полосе пропускания Δf . При этом $f_{\text{ср}} = 1$ кГц – частота среза по уровню 3 дБ; $f_2 = 2f_{\text{ср}}$ – затухание 20 дБ (рис. 2.21).

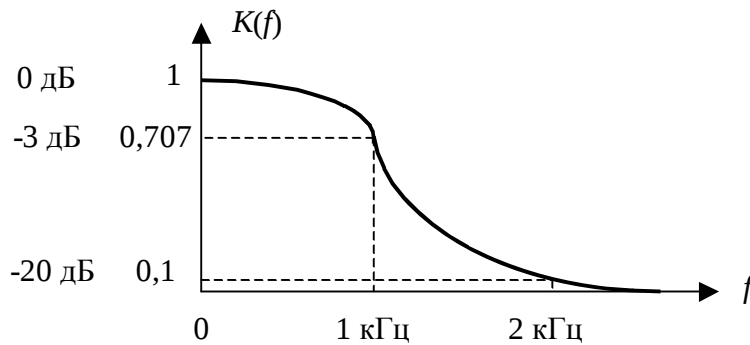


Рис. 2.21. – Требуемая АЧХ ЦФ

Выбираем для аппроксимации АЧХ аналогового фильтра-прототипа класса Баттерворта (рис. 2.22) с максимально плоской частотной характеристикой в полосе пропускания (гладкая вершина): $|K(\omega)|^2 = \frac{1}{1 + \left(\frac{\omega}{\omega_{\text{ср}}}\right)^{2n}}$,

где n – порядок фильтра Баттерворта.

1. Определяем частоты среза фильтра-прототипа:

$$\omega_a = \frac{2}{T} \operatorname{tg}\left(\frac{\omega_y T}{2}\right); T = \frac{1}{f_d} = 10^{-4} \text{ с};$$

$$\frac{f_2}{f_{\text{ср}}} = \frac{\omega_{\text{ц}2}}{\omega_{\text{ц}1}} = 2; \quad \frac{\omega_{\text{а}2}}{\omega_{\text{а}1}} = 2,24;$$

$$\omega_{\text{ц}1} = \omega_{\text{ср}}; \quad \omega_{\text{а}1} = \frac{2}{T} \operatorname{tg}\left(\frac{2\pi f_{\text{ср}} T}{2}\right) = 6498 \text{ рад/с};$$

$$\omega_{\text{ц}2} = 2\omega_{\text{ср}}; \quad \omega_{\text{а}2} = \frac{2}{T} \operatorname{tg}\left(\frac{2\pi f_2 T}{2}\right) = 14530 \text{ рад/с}.$$

2. Выбираем порядок фильтра Баттерворта n (рис. 2.22):

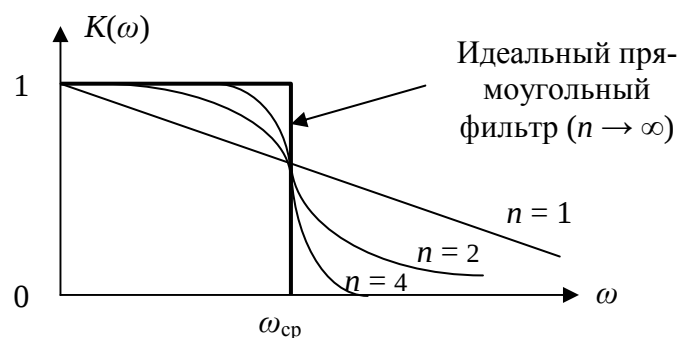


Рис. 2.22. – АЧХ фильтров Баттерворта

$$1 + \left(\frac{\omega_{a2}}{\omega_{a1}} \right)^{2n} \geq 100 \Rightarrow 1 + 2,24^{2n} \geq 100,$$

$$2n \cdot \lg(2,24) \geq 2, \text{ т.е. } n \geq \frac{1}{\lg(2,24)} = 2,85.$$

Выбираем $n = 3$. При этом затухание на $2f_{cp} \rightarrow 21$ дБ – небольшое превышение требуемого затухания.

Операторный коэффициент передачи фильтра Баттерворта имеет вид:

$$K_3(p) = \frac{1}{\left(1 + \frac{p}{\omega_{cp}} \right) \left[1 + \frac{p}{\omega_{cp}} + \left(\frac{p}{\omega_{cp}} \right)^2 \right]}, \text{ где } \omega_{cp} = \omega_{a1} = 6,5 \cdot 10^3 \text{ рад/с.}$$

Заменяем $p = \frac{2}{T} \frac{z-1}{z+1}$, после преобразования получим выражение для системной функции синтезированного ЦФ[2]:

$$H(z) = 0,018 \frac{1+z^{-1}}{1-0,5096z^{-1}} \frac{1+2z^{-1}+z^{-2}}{1-1,251z^{-1}+0,546z^{-2}}.$$

Это ЦФ третьего порядка, т.к. максимальная степень в числителе и знаменателе равна трем, что совпадает с порядком фильтра-прототипа. На рис. 2.23 приведена структурная схема синтезированного ЦФ, построенного по каскадной схеме, причем каждый из каскадов реализован в канонической форме.

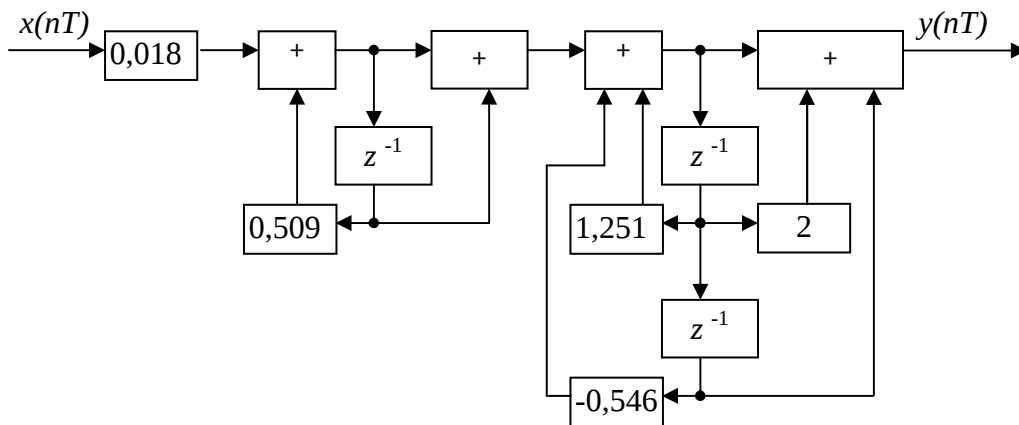


Рис. 2.23. – Синтезированный ЦФ Баттерворта 3-го порядка

2.5. Метод синтеза ЦФ с использованием Z-форм

Данный метод представляет собой дальнейшее развитие метода билинейного преобразования. Как и в предыдущем методе используется разложение $\ln(z)$ в ряд, и учитывается большее число членов ряда:

$$z = e^{pT}; \quad p = \frac{\ln(z)}{T}; \quad p^{-1} = \frac{T}{\ln(z)}; \quad \varsigma = \frac{z-1}{z+1};$$

$$p^{-1} = \frac{T}{2} \cdot \frac{1}{\varsigma + \frac{1}{3}\varsigma^3 + \frac{1}{5}\varsigma^5 + \dots + \frac{1}{2n+1}\varsigma^{2n+1}} = \frac{T}{2\varsigma} \cdot \frac{1}{1 + \left[\frac{1}{3}\varsigma^2 + \frac{1}{5}\varsigma^4 + \dots + \frac{1}{2n+1}\varsigma^{2n} \right]};$$

$$p^{-1} \cong \frac{T}{2} \cdot \frac{1}{\varsigma} = \frac{T}{2} \cdot \frac{z+1}{z-1}, \quad p^{-2} \cong \frac{T^2}{12} \cdot \frac{z^2 + 10z + 1}{(z-1)^2} \text{ и т. д.}$$

То есть для вычисления выражения для p^{-k} нужно возвести выражение для p^{-1} в степень k .

Алгоритм синтеза ЦФ заключается в следующем:

1. Находим максимальную степень в $K(p)$ (обычно степень знаменателя N больше степени числителя M). Умножаем числитель и знаменатель на p^{-N} .
2. Подставляем выражения для Z -форм (табл. 2.3).

Таблица 2.3

Z -формы операторов p^{-k} , $k = 1, 2, 3, \dots, 6$

Оператор в p -плоскости	Z -форма
p^{-1}	$\frac{T}{2} \frac{z+1}{z-1}$
p^{-2}	$\frac{T^2}{12} \frac{z^2 + 10z + 1}{(z-1)^2}$
p^{-3}	$\frac{T^3}{2} \frac{z(z+1)}{(z-1)^3}$
p^{-4}	$\frac{T^4}{6} \frac{z(z^2 + 4z + 1)}{(z-1)^4} - \frac{T^4}{720}$
p^{-5}	$\frac{T^5}{24} \frac{z(z^3 + 11z^2 + 11z + 1)}{(z-1)^5}$
p^{-6}	$\frac{T^6}{120} \frac{z(z^4 + 26z^3 + 66z^2 + 26z + 1)}{(z-1)^6} + \frac{T^6}{30240}$

3. Приводим выражение для системной функции к стандартному виду:

$$H(z) = k_0 \frac{\sum_{i=0}^M a_i z^{-i}}{1 - \sum_{j=1}^N b_j z^{-j}}.$$

4. Рассчитываем коэффициенты a_i , b_j , k_0 и строим структурную схему ЦФ по любой из форм реализации.

Преобразования частот на основе Z-форм носят существенно нелинейный характер, т. е.

$$p^{-(n+m)} = \frac{N_{n+m}(z)}{(z-1)^{n+m}} \neq \frac{N_n(z) \cdot N_m(z)}{(z-1)^{n+m}}.$$

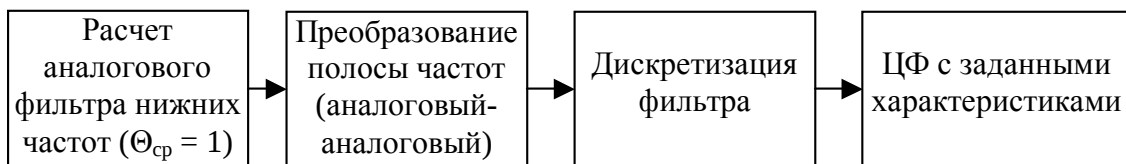
Достоинства метода:

1. Не нужно производить трансформацию частот аналогового фильтра-прототипа.
2. Чем выше порядок аналогового фильтра-прототипа, тем точнее передаются характерные точки АЧХ аналогового фильтра-прототипа.

2.6. Частотные преобразования, применяемые при синтезе цифровых фильтров

Основные методы частотных преобразований, применяемых при синтезе ЦФ показаны на рис. 2.24. В первом методе происходит преобразование полосы частот аналогового фильтра, а во втором – преобразование полосы частот цифрового фильтра. Метод частотных преобразований ЦФ (метод 2) предложил американский ученый А. Дж. Константи́нидис. По данному методу рассчитывается низкочастотный цифровой фильтр, а затем он преобразуется в фильтры нижних и верхних частот, полосовые и режекторные фильтры (ФНЧ, ФВЧ, ПФ и РФ, соответственно).

Метод 1



Метод 2

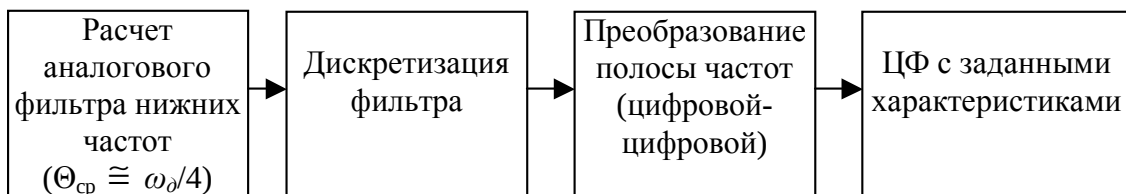


Рис. 2.24. – Основные методы частотных преобразований

При использовании первого метода для преобразования полосы частот делаются замены оператора p , по соотношениям, приведенным в табл. 2.4.

Таблица 2.4

Частотные преобразования аналогового фильтра

1.	ФНЧ $\rightarrow$ ФНЧ ₁	$p \rightarrow \frac{p}{\omega_{cp}}$
2.	ФНЧ $\rightarrow$ ФВЧ	$p \rightarrow \frac{\omega_{cp}}{p}$
3.	ФНЧ $\rightarrow$ ПФ	$p \rightarrow \frac{p^2 + \omega_1 \omega_2}{p(\omega_2 - \omega_1)}$
4.	ФНЧ $\rightarrow$ РФ	$p \rightarrow \frac{p(\omega_2 - \omega_1)}{p^2 + \omega_1 \omega_2}$

Здесь ω_{cp} – частота среза ФНЧ₁, ФВЧ, а ω_1 , ω_2 – нижняя и верхняя частоты среза ПФ и РФ.

Пример. Возьмем в качестве нормализованного ФНЧ фильтр Баттерворта второго порядка:

$$K(p) = \frac{\Theta_{cp}^2}{p^2 + \sqrt{2}\Theta_{cp}p + \Theta_{cp}^2} \Big|_{\Theta_{cp}=1} = \frac{1}{p^2 + \sqrt{2}p + 1}.$$

Делая замену $p \rightarrow \frac{p}{\omega_{cp}}$ (из таблицы 2.4), получим ФНЧ₁:

$$K_1(p) = \frac{\omega_{cp}^2}{p^2 + \sqrt{2}\omega_{cp}p + \omega_{cp}^2}, \text{ где } \omega_{cp} \text{ – круговая частота среза ФНЧ}_1.$$

Производя аналогичные замены из таблицы 2.4, получаем ФВЧ, ПФ, РФ.

Общие частотные преобразования ЦФ по Константи́нидису приведены в табл. 2.5.

Таблица 2.5

Общие частотные преобразования ЦФ по Константи́нидису

Тип фильтра	Преобразование	Расчетные формулы
1. ФНЧ ₁	$z^{-1} \rightarrow \frac{z^{-1} - \alpha}{1 - z^{-1}\alpha}$	$\alpha = \frac{\sin\left[\frac{T}{2}(\Theta_{cp} - \omega_{cp})\right]}{\sin\left[\frac{T}{2}(\Theta_{cp} + \omega_{cp})\right]},$ где ω_{cp} – требуемая частота среза ФНЧ₁, Θ_{cp} – частота среза исходного фильтра.

Таблица 2.5 (продолжение)

2. ФВЧ	$z^{-1} \rightarrow -\frac{z^{-1} + \alpha}{1 + z^{-1}\alpha}$	$\alpha = -\frac{\cos\left[\frac{T}{2}(\omega_{cp} + \Theta_{cp})\right]}{\cos\left[\frac{T}{2}(\omega_{cp} - \Theta_{cp})\right]},$ где ω_{cp} – требуемая частота среза ФВЧ, Θ_{cp} – частота среза исходного фильтра.
3. ПФ	$z^{-1} \rightarrow -\frac{z^{-2} - \frac{2\alpha \cdot k}{k+1}z^{-1} + \frac{k-1}{k+1}}{\frac{k-1}{k+1}z^{-2} - \frac{2\alpha \cdot k}{k+1}z^{-1} + 1}$	$\alpha = \frac{\cos\left[\frac{T}{2}(\omega_2 + \omega_1)\right]}{\cos\left[\frac{T}{2}(\omega_2 - \omega_1)\right]} = \cos(\omega_0 T)$, $k = \operatorname{ctg}\left[\frac{T}{2}(\omega_2 - \omega_1)\right] \cdot \operatorname{tg}\left(\frac{\Theta_{cp} T}{2}\right)$, где ω_1, ω_2 – нижняя и верхняя частоты среза ПФ, Θ_{cp} – частота среза исходного ФНЧ, ω_0 – центральная частота ПФ.
4. РФ	$z^{-1} \rightarrow -\frac{z^{-2} - \frac{2\alpha}{1+k}z^{-1} + \frac{1-k}{1+k}}{\frac{1-k}{1+k}z^{-2} - \frac{2\alpha}{1+k}z^{-1} + 1}$	$\alpha = \frac{\cos\left[\frac{T}{2}(\omega_2 + \omega_1)\right]}{\cos\left[\frac{T}{2}(\omega_2 - \omega_1)\right]} = \cos(\omega_0 T)$, $k = \operatorname{tg}\left[\frac{T}{2}(\omega_2 - \omega_1)\right] \operatorname{tg}\left(\frac{T}{2}\Theta_{cp}\right),$ где ω_1, ω_2 – нижняя и верхняя частоты среза РФ, Θ_{cp} – частота среза исходного ФНЧ, ω_0 – центральная частота РФ.

На рис. 2.25 приведены амплитудно-частотные характеристики исходного фильтра (ФНЧ) и преобразованных фильтров (ФНЧ1, ФВЧ, ПФ и РФ). В качестве исходного ФНЧ взят фильтр Чебышева пятого порядка с пульсациями АЧХ в полосе пропускания и $\Theta_{cp} = 1$. Угловая частота дискретизации

$$\omega_0 = \frac{2\pi}{T} = 6,28, \quad T = 1 \text{ сек.}$$

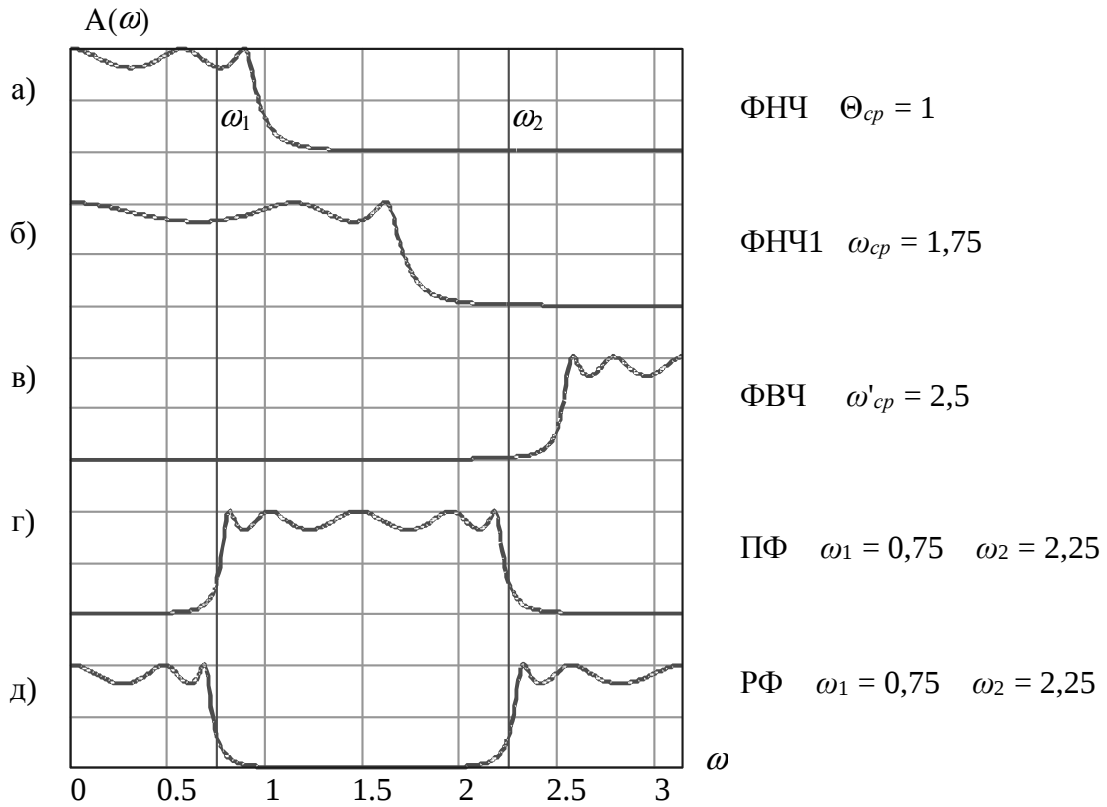


Рис. 2.25. – АЧХ исходного (ФНЧ) и преобразованных (ФНЧ1, ФВЧ, ПФ, РФ) фильтров

При расчете полосового и режекторного фильтров справедливо соотношение:

$$\alpha = \cos(\omega_0 T) = \frac{\cos\left[\frac{T}{2}(\omega_2 + \omega_1)\right]}{\cos\left[\frac{T}{2}(\omega_2 - \omega_1)\right]}, \quad 0 < \omega_1 < \omega_2 < \frac{\omega_\delta}{2}.$$

Это означает, что из трех параметров ω_0 , ω_1 , ω_2 только два независимы, а третий является их функцией. Чаще всего в качестве независимых параметров выбирают нижнюю ω_1 и верхнюю ω_2 частоты полосы пропускания (режекции) ЦФ.

$$\text{Тогда } \omega_0 = \frac{1}{T} \arccos(\alpha) = \frac{1}{T} \arccos\left(\frac{\cos\left[\frac{T}{2}(\omega_2 + \omega_1)\right]}{\cos\left[\frac{T}{2}(\omega_2 - \omega_1)\right]}\right).$$

Если $\Delta\omega = \omega_2 - \omega_1 \ll \frac{\omega_2 + \omega_1}{2}$, а $\omega_1 \neq 0$ и $\omega_2 \neq \frac{\omega_\delta}{2}$,

$$\text{то } \cos\left[\frac{T}{2}(\omega_2 - \omega_1)\right] \rightarrow 1 \text{ и } \omega_0 \approx \frac{1}{T} \arccos\left(\cos\left[\frac{T}{2}(\omega_2 + \omega_1)\right]\right) = \frac{\omega_2 + \omega_1}{2}.$$

При $\omega_1 = 0$ получим $\omega_0 = \frac{1}{T} \arccos(1) = 0$.

При $\omega_2 = \frac{\omega_\delta}{2} = \frac{\pi}{T}$ отношение косинусов $\frac{\cos\left[\frac{T}{2}(\omega_2 + \omega_1)\right]}{\cos\left[\frac{T}{2}(\omega_2 - \omega_1)\right]} = -1$ и

$$\omega_0 = \frac{1}{T} \arccos(-1) = \frac{\pi}{T} = \frac{\omega_\delta}{2}.$$

Таким образом, значения ω_1 и ω_2 выбираются из следующего условия:
 $0 < \omega_1 < \omega_2 < \frac{\omega_\delta}{2} = \frac{\pi}{T}$, а центральная частота полосы пропускания (режекции) оценивается как $\omega_0 \cong \frac{\omega_2 + \omega_1}{2}$.

Обычно задача преобразования ЦФ ставится следующим образом: задан ЦФ нижних частот (ФНЧ-прототип) с определенной частотой среза, например Θ_{cp} , из которого путем указанного преобразования (табл. 2.5) можно получить ФНЧ1, ФВЧ, ПФ, РФ. Исходный ФНЧ-прототип, как правило, нормируется по частоте к величине, равной $1/4$ частоты дискретизации ω_δ , т.е. $\Theta_{cp} = \frac{1}{4} \omega_\delta = \frac{\pi}{2} T$. Такой фильтр называется *нормализованным*.

Примечание. Нормализованный цифровой ФНЧ легко получить из аналогового ФНЧ с частотой среза $\omega_{cp} = \frac{2}{T}$ методом БЛП. По закону трансформации частот частота среза ЦФ будет равной:

$$\Theta_{cp} = \frac{2}{T} \arctg\left(\frac{\omega_{cp} T}{2}\right) = \frac{2}{T} \arctg(1) = \frac{2}{T} \cdot \frac{\pi}{4} = \frac{\omega_\delta}{4}.$$

Синтез нормализованного фильтра методом БЛП существенно упрощается из-за сокращения множителей ω_{cp} и $2/T$.

Нормализованные ЦФ наиболее удобно использовать при частотных преобразованиях, особенно при синтезе полосовых и режекторных фильтров с переменной полосой пропускания (режекции).

Рассмотрим два важных частных случая синтеза цифрового полосового фильтра:

- 1) $\alpha \neq 0$, $k = 1$,
- 2) $\alpha = 0$, $k \neq 1$.

В первом случае из условия $k = 1$ следует, что полоса пропускания ЦФ постоянна и равна $\Delta\omega = \omega_2 - \omega_1 = \Theta_{cp}$, а центральная частота управляется

параметром α , т. е. $\omega_0 = \frac{1}{T} \arccos(\alpha)$. При $\alpha \in (-1, 1)$, $\omega_0 \in \left(0, \frac{\omega_\delta}{2}\right)$. Причем,

при $\alpha = -1$, $\omega_0 = \frac{\omega_\delta}{2}$; при $\alpha = 0$, $\omega_0 = \frac{\omega_\delta}{4}$; при $\alpha = 1$, $\omega_0 = 0$.

Во втором случае из условия $\alpha = 0$ следует, что центральная частота постоянна и равна $\omega_0 = \frac{1}{T} \arccos(0) = \frac{1}{T} \cdot \frac{\pi}{2} = \frac{\omega_\delta}{4}$ (используем нормализованный фильтр), а полоса пропускания управляется параметром $k \in (0, \infty)$. Причем при $k = 0$, $\Delta\omega = \omega_2 - \omega_1 = \frac{\omega_\delta}{2}$;

$$k = 1, \quad \Delta\omega = \omega_2 - \omega_1 = \Theta_{cp}, \quad 0 < \Theta_{cp} < \frac{\omega_\delta}{2};$$

$$k = \infty, \quad \Delta\omega = \omega_2 - \omega_1 = 0.$$

Предельные значения параметров α и k являются вырожденными случаями и при частотных преобразованиях по Константи́нидису не используются.

2.7. Прямой синтез ЦФ

Прямой синтез ЦФ сводится к нахождению функции, аппроксимирующей заданную частотную характеристику так, что бы в результате аппроксимации получилось дробно-рациональное выражение для $H(e^{j\omega T})$ (производим замену оператора p на функции $\sin(\omega T)$, $\cos(\omega T)$, $\operatorname{tg}(\omega T/2)$, $\operatorname{ctg}(\omega T/2)$, которые затем преобразуем в $e^{\pm j\omega T}$) [2].

Для перехода от $H(e^{j\omega T})$ к системной функции $H(z)$ необходимо в выражении для КЧХ $e^{j\omega T}$ заменить на z . Чтобы выражение для системной функции получилось дробно-рациональным, необходимо, чтобы исходное выражение было также дробно-рациональным относительно $e^{j\omega T}$. Поиск подходящей аппроксимации для КЧХ является главной трудностью метода прямого синтеза ЦФ.

Обычно задачу аппроксимации решают, подбирая для $K(j\omega)$ дробно-рациональное выражение относительно тригонометрических функций от ωT . Каждая из тригонометрических функций может быть представлена в виде комбинации комплексных экспонент $e^{j\omega T}$ и $e^{-j\omega T}$, в результате чего можно получить выражение для частотной характеристики $K(j\omega)$ в требуемой форме. Делая замену $e^{j\omega T}$ на z , выполняем переход от $K(j\omega)$ к $H(z)$. Для этого тригонометрические функции представляем в виде:

$$\sin(\omega T) = \frac{1}{2j} [e^{j\omega T} - e^{-j\omega T}],$$

$$\cos(\omega T) = \frac{1}{2} [e^{j\omega T} + e^{-j\omega T}],$$

$$tg(\omega T) = \frac{\sin(\omega T)}{\cos(\omega T)} \text{ и т.д.}$$

Например, для фильтра нижних частот с частотной среза ω_{cp} , частотная характеристика может быть аппроксимирована выражением

$$|K(j\omega)|^2 = \frac{1}{1 + \left[ctg\left(\frac{\omega_{cp}T}{2}\right) \cdot tg\left(\frac{\omega T}{2}\right) \right]^{2n}}, \quad \omega \in \left[0, \frac{2\pi}{T}\right].$$

Графики АЧХ такого фильтра близки к АЧХ фильтра Баттерворта, приведенным на рис. 2.22. Здесь $n = 1, 2, 3, \dots$ – порядок фильтра НЧ.

Процедура синтеза ЦФ данным методом сводится к следующему:

1. По заданным характеристикам находим порядок фильтра n (как и в примере для билинейного преобразования).
2. Заменяем в выражении для $|K(j\omega)|^2 e^{j\omega T}$ на z , затем находим выражение для квадрата системной функции $|H(z)|^2 = H(z)H(z^{-1})$.
3. Находим положение полюсов $|H(z)|^2$, а затем системной функции $H(z)$ (они должны лежать внутри единичной окружности $|z| < 1$).
4. Исходя из расположения полюсов, находим выражение для системной функции $H(z)$ и записываем разностное уравнение, на основании которого построим фильтр (его структурную схему).

Более подробно прямые методы синтеза ЦФ изложены в литературе [2, 5].

2.8. Методы синтеза фильтров с КИХ

До сих пор мы синтезировали ЦФ с бесконечной импульсной характеристикой (БИХ). Однако наряду с ними широкое распространение получили ЦФ с ИХ конечной длительности или КИХ-фильтры.

Главное достоинство КИХ-фильтров в том, что они всегда устойчивы.

КИХ-фильтры не имеют непосредственных аналогов среди пассивных цепей (т.е. нет фильтров-прототипов), поэтому методы их синтеза относятся к прямым методам синтеза ЦФ.

Одним из наиболее часто применяемых методов синтеза КИХ-фильтров является метод частотной выборки.

2.8.1. Метод частотной выборки

Идея этого метода довольно проста. Амплитудно-частотную характеристику аналогового фильтра дискретизируют как периодическую функцию, разбивая интервал $\omega_i \in \left[-\frac{\pi}{T} < \omega < \frac{\pi}{T}\right]$ на N равных частей (N – нечетное). Ин-

тервал дискретизации $\Omega = \frac{\omega_d}{N} = \frac{2\pi}{NT}$ должно быть таким, что бы передать все особенности АЧХ. К полученной последовательности значений $K(n\Omega)$ применяют обратное дискретное преобразование Фурье (ОДПФ) и находят импульсную характеристику $h(nT)$. Значения $h(nT)$ являются коэффициентами нерекурсивного фильтра.

Число выборок в АЧХ соответствует числу отсчетов ИХ.

Найденная ИХ соответствует физически нереализуемому фильтру, поэтому ее сдвигают вправо на половину своей длины (в этом случае к каждому аргументу nT в выражении для ИХ добавится отрицательная величина, равная $(N-1)T/2$). При таком смещении изменится ФЧХ, а АЧХ останется неизменной, т.е. можно сказать, что метод частотной выборки позволяет синтезировать фильтры с заданной АЧХ при неконтролируемой ФЧХ.

После нахождения ИХ можно записать РУ:

$$y(nT) = \sum_{i=0}^N a_i x(nT - iT), \quad a_i = h(iT), \quad n \geq 0, \quad \text{на основании которого можно}$$

построить фильтр.

Пример. Пусть требуется синтезировать цифровой фильтр с АЧХ $A(\omega)$ по заданной АЧХ аналогового фильтра-прототипа $K(\omega) = e^{-\left(\frac{\omega}{\omega_0}\right)^2}$, где $\omega_0 = 2\pi \cdot 200 \text{ рад/с}$.

Продискретизируем АЧХ на 5 отсчетов ($N = 5$) с интервалом дискретизации $\Omega = \omega_0$. Тогда $K(0) = e^0 = 1$, $K(\pm\Omega) = e^{-1} = 0,368$, $K(\pm 2\Omega) = e^{-4} = 0,0183$ (рис. 2.26).

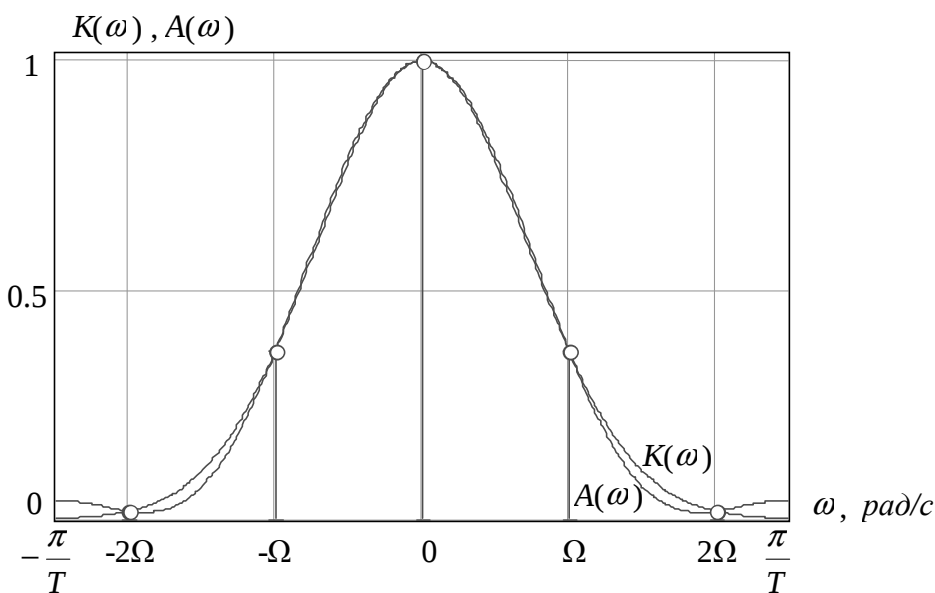


Рис. 2.26. – Дискретизация гауссоидальной $K(\omega)$ и АЧХ ЦФ – $A(\omega)$

Применяя к $K(n\Omega)$ обратное дискретное преобразование Фурье, найдем выражение для импульсной характеристики (ИХ):

$$h(nT) = \frac{1}{N} \sum_{k=-\frac{N-1}{2}}^{\frac{N-1}{2}} K(k\Omega) e^{jkn\frac{2\pi}{N}}, \quad n \in \left[-\frac{N-1}{2}, \frac{N-1}{2} \right].$$

В нашем случае $N = 5$ (N – число степеней свободы сигнала):

$$h(nT) = \frac{1}{5} \sum_{k=-2}^2 K(k\Omega) e^{jkn\frac{2\pi}{5}}, \quad n \in [-2, 2].$$

Вычислим значение ИХ при $n = 0$:

$$h(0) = \frac{1}{5} \sum_{k=-2}^2 K(k\Omega) e^{jk \cdot 0 \cdot \frac{2\pi}{5}} = \frac{1 + 2 \cdot 0,368 + 2 \cdot 0,0183}{5} = 0,355;$$

$$h(\pm T) = 0,239; \quad h(\pm 2T) = 0,083.$$

$$\text{ИХ: } h(nT) = 0,083h(nT + 2T) + 0,239h(nT + T) + 0,355h(nT) + 0,239h(nT - T) + 0,083h(nT - 2T), \quad -\infty < n < \infty.$$

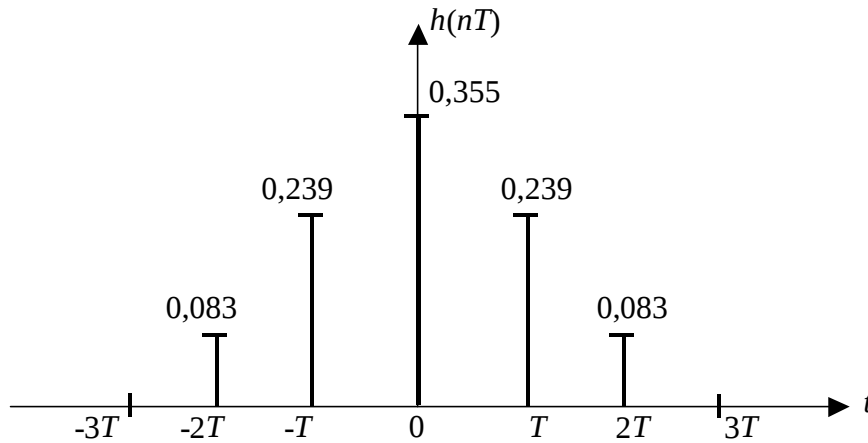


Рис. 2.27. – ИХ физически нереализуемого КИХ-фильтра

Для того чтобы фильтр стал физически реализуемым, сдвинем отсчеты ИХ вправо на величину $\frac{N-1}{2}$ (N – нечетное). Запишем разностное уравнение:

$$y(nT) = 0,083x(nT) + 0,239x(nT - T) + 0,355x(nT - 2T) + 0,239x(nT - 3T) + 0,083x(nT - 4T), \quad n \geq 0.$$

Таким образом, мы получим нерекурсивный КИХ-фильтр 4-го порядка (рис. 2.28).

Коэффициенты фильтра a_i можно пронормировать к коэффициенту $a_2 = 0,355$. Вид АЧХ фильтра не изменится при существенном упрощении реализации. Нормированные коэффициенты примут значения:

$$a_0 = a_4 = 0,234; \quad a_1 = a_3 = 0,673; \quad a_2 = 1.$$

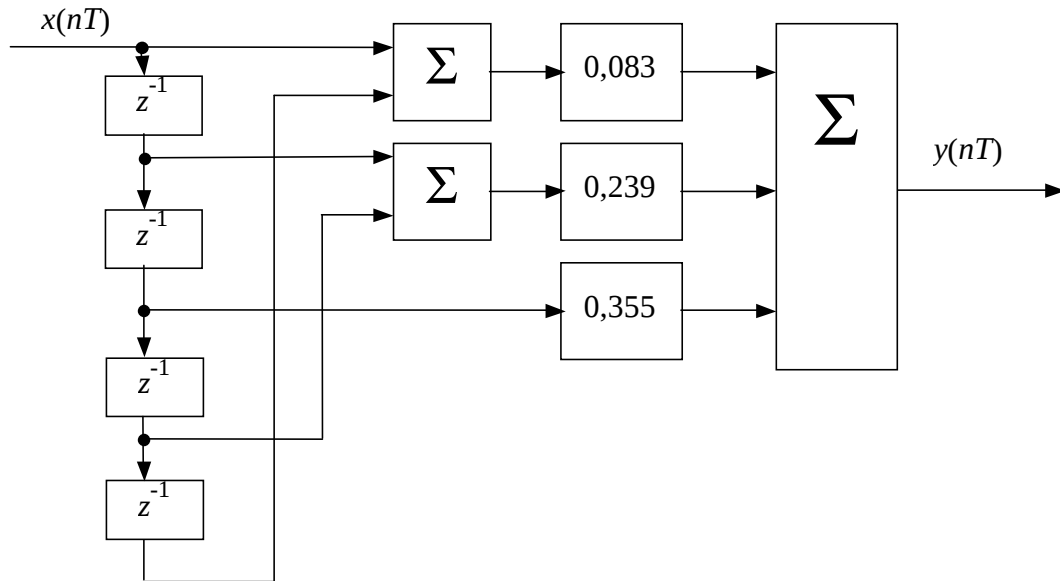


Рис. 2.28. – Структурная схема НРЦФ

Амплитудно-частотная характеристика $A(\omega)$ синтезированного цифрового фильтра мало отличается от исходной $K(\omega)$ на интервале частот $\omega_u \in \left[-\frac{\pi}{T}, \frac{\pi}{T}\right]$ (рис. 2.26). Ошибка составляет не более 4 %. Такой хороший результат получается при малом числе отсчетов АЧХ $N = 5$ благодаря тому, что исходная АЧХ гауссова фильтра имеет плавный вид.

2.8.2. Метод временных окон

Предыдущий метод хорошо подходит для фильтров с плавной АЧХ. Для фильтров, имеющих АЧХ с крутыми склонами (большим затуханием) процедура синтеза существенно усложняется.

Возьмем к примеру ФНЧ с идеальной АЧХ (прямоугольной) и подвергнем ее дискретизации. Получившаяся АЧХ точно пройдет через все отсчетные точки, но в промежутках между ними будут большие колебания. Увеличение числа отсчетов не помогает – увеличивается изрезанность АЧХ, а амплитуда паразитных колебаний не изменяется и составляет примерно 14 % от размаха АЧХ в точке разрыва (рис. 2.29).

Уровень боковых лепестков в полосе затухания остается также высоким. Данный эффект называется явлением Гиббса. Впервые данный эффект был изучен в связи с усечением ряда Фурье, применяемого для разложения по гармоникам периодических сигналов, в том числе и периодической АЧХ, характерной для цифровых фильтров.

Применение метода временных окон позволяет уменьшить или даже устранить изрезанность АЧХ (т.е. исчезнут колебания на ней), однако при этом склоны полученной АЧХ будут более пологими, чем у исходной АЧХ.

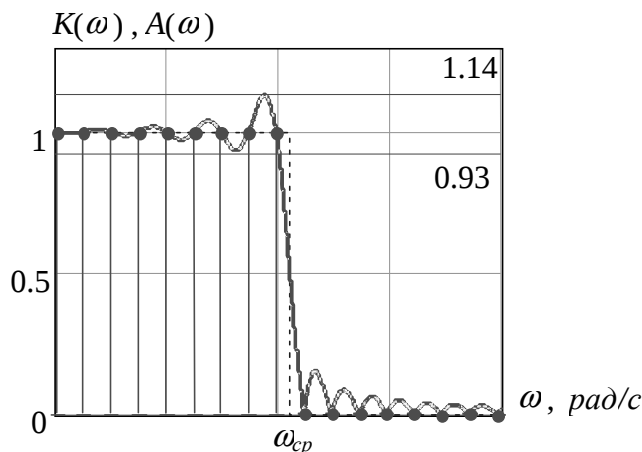


Рис. 2.29. – Фильтр с КИХ, построенный по методу частотной выборки

Пример. Пусть АЧХ идеального ФНЧ с частотой среза ω_{cp} описывается выражением:

$$K(\omega) = \begin{cases} 1, & \text{если } -\omega_{cp} \leq \omega \leq \omega_{cp}, \\ 0, & \text{иначе.} \end{cases}$$

Импульсная характеристика такого фильтра описывается функцией:

$$h_a(nT) = \frac{\omega_{cp}}{\pi} \cdot \frac{\sin(\omega_{cp}t)}{\omega_{cp}t}, \quad -\infty < t < \infty.$$

Из данной импульсной характеристики можно сделать бесконечную выборку значений. Интервал дискретизации делают достаточно малым, чтобы передать все особенности ИХ. Обычно руководствуются правилом $\omega_\delta \geq 4\omega_{cp}$, $T = \frac{2\pi}{\omega_\delta}$.

Усечение продискретизированной ИХ может быть произведено на любом уровне N , однако это усечение приводит к сильной изрезанности АЧХ в полосе пропускания и большому уровню боковых лепестков. Обычно число отсчетов выбирают достаточно большим: $N = 21; 51; 101; 201; 501$ и т. д.

Для обеспечения физической реализуемости импульсную характеристику сдвигают на половину своей длительности, т.е. $\frac{N-1}{2}$. Сдвиг импульсной характеристики не изменяет вида АЧХ, меняется только фазочастотная характеристика. На рис. 2.30а усеченная и сдвинутая ИХ показана при $N = 81$.

Для уменьшения изрезанности АЧХ отсчеты ИХ умножают на отсчеты некоторой весовой функции $w(nT)$ (рис. 2.30б), называемой временным окном. Временное окно обычно симметричная функция, медленно спадающая к краям. Центральный отсчет временного окна нормируют к единице. В результате получаем измененную ИХ вида $h(nT)w(nT)$ (рис. 2.30в).

При использовании временного окна существенно уменьшается изрезанность АЧХ и уровень боковых лепестков синтезированного по методу частотной выборки ЦФ, но платой за это будет более пологий склон АЧХ синтезированного таким образом цифрового фильтра (рис. 2.30д).

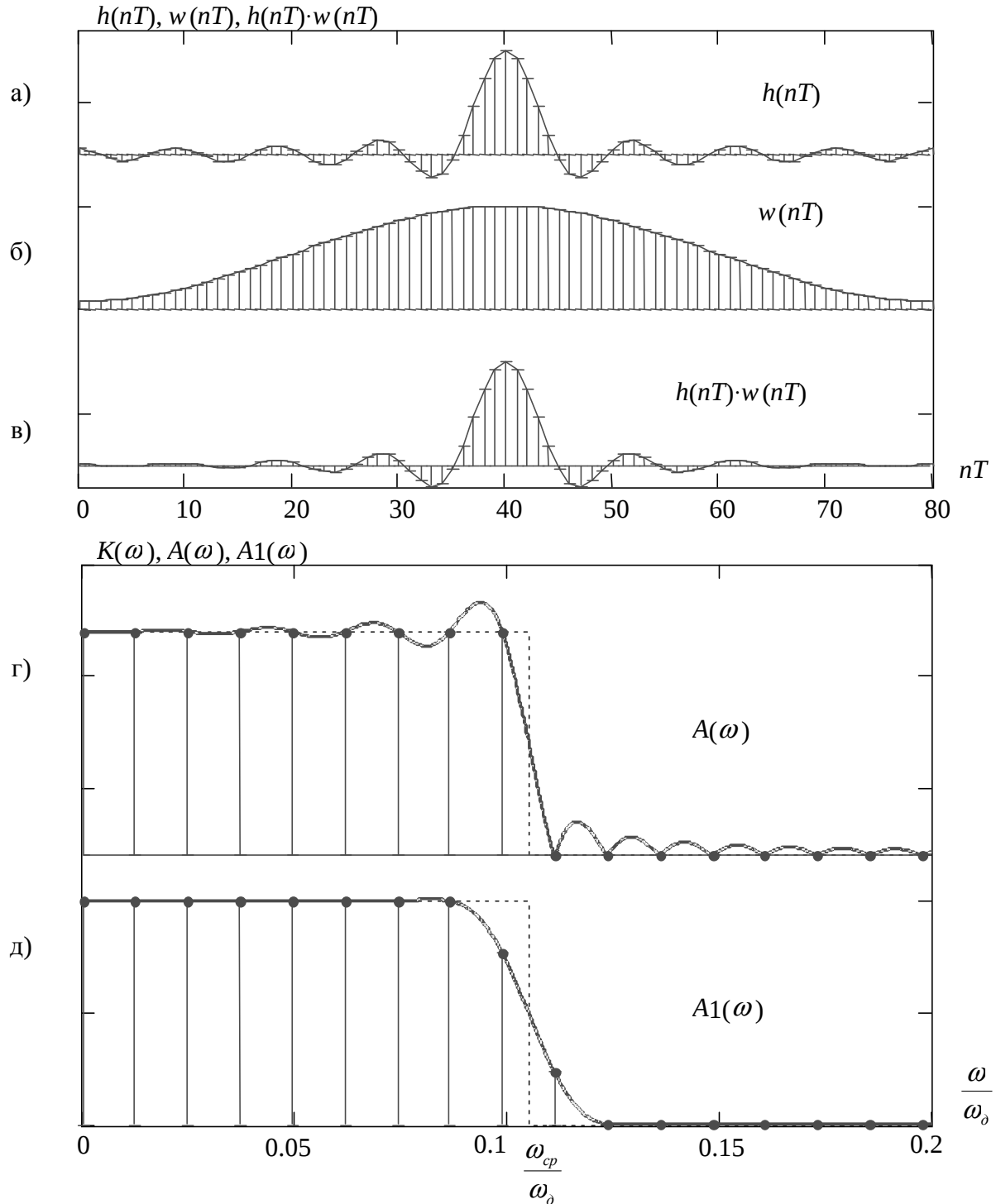


Рис. 2.30. – Иллюстрация метода "временных окон"

Временное окно представляет собой функцию определенного вида. Рассмотрим основные используемые временные окна:

1. Окно Дирихле (прямоугольное окно):

$$w(n) = \begin{cases} 1, & 0 \leq n \leq N-1 \\ 0, & N-1 < n < 0 \end{cases}. \text{ Применяют для ограничения длительности ИХ.}$$

2. Окно Хемминга:

$$w(n) = \begin{cases} 0,54 - 0,46 \cos\left(\frac{2\pi n}{N-1}\right), & 0 \leq n \leq N-1 \\ 0, & N-1 < n < 0 \end{cases}.$$

Имеет малый уровень боковых лепестков -53,6 дБ. Частотная характеристика близка к прямоугольной.

3. Окно Бартлетта (треугольное окно):

$$w(n) = \begin{cases} \frac{2n}{N-1}, & 0 \leq n \leq \frac{N-1}{2} \\ 2 - \frac{2n}{N-1}, & \frac{N-1}{2} < n \leq N-1 \\ 0, & N-1 < n < 0 \end{cases}.$$

4. Окно Ханна:

$$w(n) = \begin{cases} \frac{1}{2} \left[1 - \cos\left(\frac{2\pi n}{N-1}\right) \right], & 0 \leq n \leq N-1 \\ 0, & N-1 < n < 0 \end{cases}.$$

5. Окно Блэкмана:

$$w(n) = \begin{cases} 0,42 - 0,5 \cos\left(\frac{2\pi n}{N-1}\right) + 0,08 \cos\left(\frac{4\pi n}{N-1}\right), & 0 \leq n \leq N-1 \\ 0, & N-1 < n < 0 \end{cases}.$$

6. Окно Кайзера (обобщенное временное окно):

$$w(n) = \begin{cases} \frac{I_0 \left[\omega_a \sqrt{\left(\frac{N-1}{2}\right)^2 - \left(n - \frac{N-1}{2}\right)^2} \right]}{I_0 \left[\omega_a \left(\frac{N-1}{2}\right) \right]}, & 0 \leq n \leq N-1, \\ 0 & \text{иначе.} \end{cases}$$

Окно Кайзера наиболее близкое к оптимальному.

Здесь $I_0[\cdot]$ – модифицированная функция Бесселя нулевого порядка, параметр ω_a применяют для настройки окна Кайзера. При уменьшении ω_a уменьшается уровень боковых лепестков АЧХ и пульсаций в полосе пропускания, но также уменьшается и крутизна склонов АЧХ. При увеличении этого параметра можно достичь формы АЧХ близкой к прямоугольной (идеальной)

при допустимом уровне боковых лепестков и пульсаций в полосе пропускания. Типичные значения $\omega_a \frac{N-1}{2}$ лежат в пределах от 4 до 9. Модифицированную функцию Бесселя нулевого порядка рассчитывают как сумму ряда, ограничиваясь двадцатью членами ряда [5]: $I_0(x) = 1 + \sum_{k=0}^{\infty} \left[\frac{1}{k!} \cdot \left(\frac{x}{2} \right)^k \right]^2$.

Примеры некоторых окон приведены на рис. 2.31.

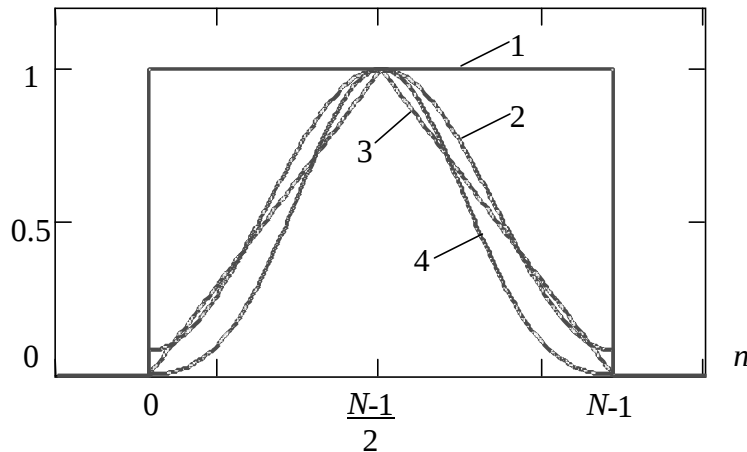


Рис. 2.31. – Временные окна:
 1 – окно Дирихле; 2 – окно Хемминга;
 3 – окно Бартлетта; 4 – окно Блэкмана.

В настоящее время используется более 50 видов временных окон.

Алгоритм синтеза по методу временных окон:

1. Задаемся требуемой АЧХ – ФНЧ, ФВЧ, ПФ, РФ.

2. Дискретизируем АЧХ на N частей на интервале частот $\omega_y \in \left[-\frac{\pi}{T}, \frac{\pi}{T} \right]$.

3. К полученным отсчетам АЧХ применяем обратное дискретное преобразование Фурье (ОДПФ) и получаем отсчеты ИХ $h_\delta(nT)$.

4. Сдвигаем отсчеты ИХ вправо на величину $\frac{N-1}{2}$ для достижения физической реализуемости ЦФ.

5. Используя одно из временных окон, получаем отсчеты ИХ $h(nT) = h_\delta(nT) \cdot w(nT)$, $0 \leq n \leq N-1$.

6. Значения отсчетов ИХ принимаем за коэффициенты НРЦФ: $a_n = h(nT)$, где $n = \overline{0, N-1}$.

7. Записываем выражение для системной функции: $H(z) = \sum_{i=0}^{N-1} a_i z^{-i}$.

8. Контролируем АЧХ ЦФ: $A(\omega) = |H(e^{j\omega T})|$.

9. Программируем ЦФ, либо реализуем его аппаратным способом.

Далее, на листинге, приведен пример программной реализации цифрового фильтра нижних частот с помощью пакета MathCAD.

**Пример программирования цифрового фильтра с использованием окна Хэмминга
в системе MCAD 6.0 (листинг программы)**

Требуется синтезировать ЦФ нижних частот (за основу взят ПФ с $f_1 = 0$ кГц):
 $f_1 = 0$ кГц, $f_2 = 3$ кГц, $f_d = 10$ кГц, $N = 81$

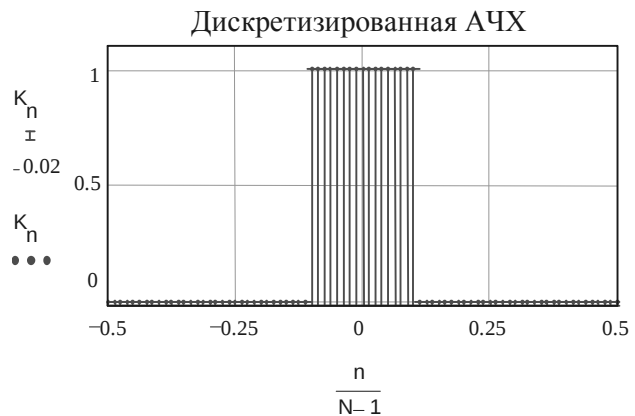
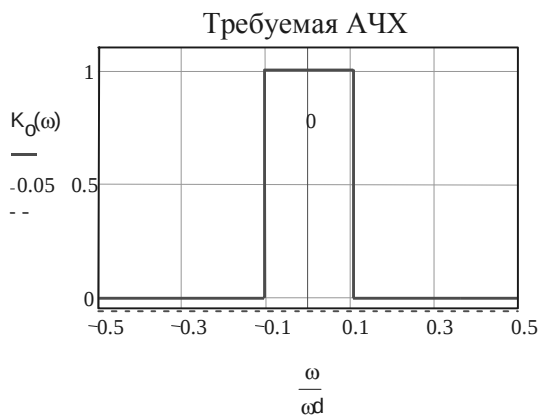
kHz := $10^3 \cdot \text{sec}^{-1}$ f1 := 0·kHz f2 := 1.05·kHz fd := 10·kHz ORIGIN := -100
T := $10^{-4} \cdot \text{sec}$ $\omega_1 := 2 \cdot \pi \cdot f_1$ $\omega_2 := 2 \cdot \pi \cdot f_2$ $\omega_d := 2 \cdot \pi \cdot f_d$ N := 81

1. Дискретизируем частотную характеристику ПФ $\omega \in [-\pi/T, \pi/T]$, $N = 81$:

$$K_0(\omega) := \Phi(\omega + \omega_2) - \Phi(\omega + \omega_1) + \Phi(\omega - \omega_1) - \Phi(\omega - \omega_2)$$

$$\Delta\omega := \frac{\omega_d}{N} \quad n := -\frac{N-1}{2} \dots \frac{N-1}{2} \quad K_n := K_0(n \cdot \Delta\omega)$$

$$\text{step} := .0005 \cdot \omega_d \quad \omega := -.5 \cdot \omega_d, -.5 \cdot \omega_d + \text{step} \dots .5 \cdot \omega_d \quad N = 81$$



2. Получаем с помощью ОДПФ ИХ цифрового физически нереализуемого фильтра:

$$hd(n) := \frac{1}{N} \cdot \sum_{k=-\frac{N-1}{2}}^{\frac{N-1}{2}} K_k \cdot \exp\left(j \cdot n \cdot k \cdot \frac{2 \cdot \pi}{N}\right)$$

$$i := -0.2 \cdot (N-1) \dots 0.2 \cdot (N-1) \quad j := -0.2 \cdot (N-1), -0.19 \cdot (N-1) \dots 0.2 \cdot (N-1)$$

3. Делаем фильтр физически реализуемым [сдвигаем ИХ на $0.5 \cdot (N-1)$]:

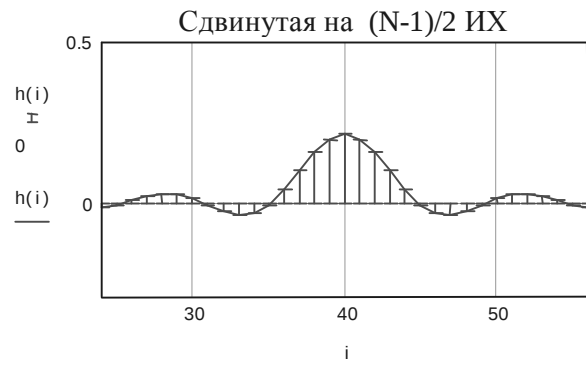
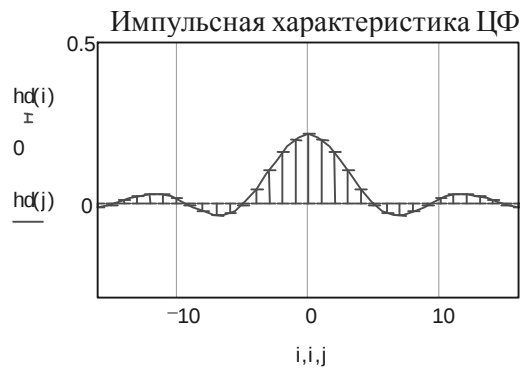
$$n := 0 \dots N-1 \quad h(n) := hd\left(n - \frac{N-1}{2}\right) \quad i := .3 \cdot (N-1) \dots .7 \cdot (N-1)$$

4. Определяем окно Хэмминга:

$$w(n) := \text{if}\left(n < 0, 0, \text{if}\left(n > N-1, 0, 0.54 - 0.46 \cdot \cos\left(\frac{2 \cdot \pi \cdot n}{N-1}\right)\right)\right)$$

5. Находим коэффициенты нерекурсивного фильтра:

$$i := 0 \dots N-1 \quad a1_i := h(i) \quad a2_i := h(i) \cdot w(i)$$



6. Записываем выражение для системной функции:

$$H1(z) := \sum_i a1_i \cdot z^{-i}$$

$$H2(z) := \sum_i a2_i \cdot z^{-i}$$

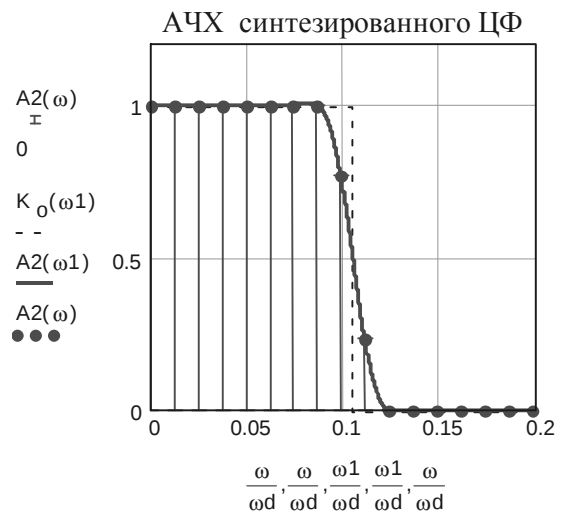
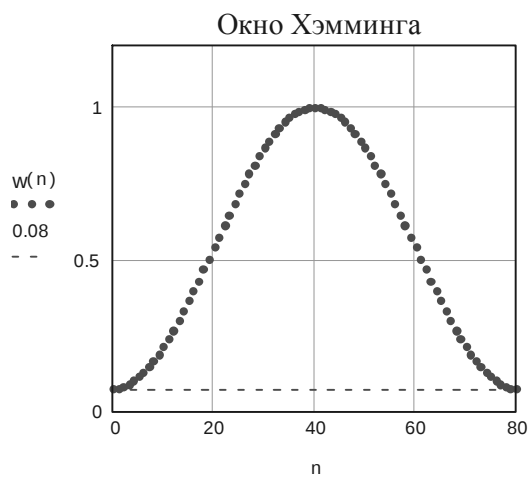
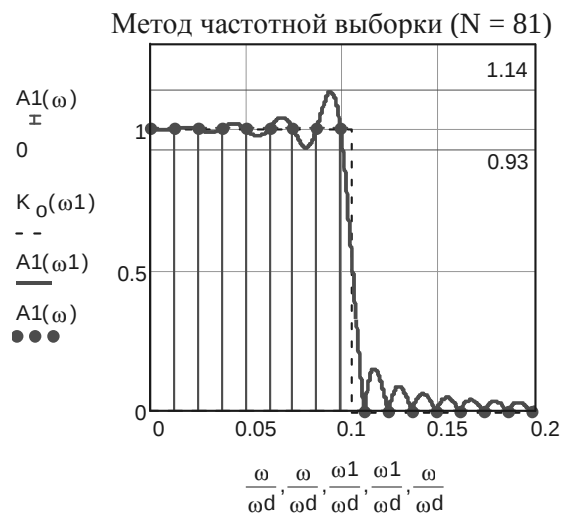
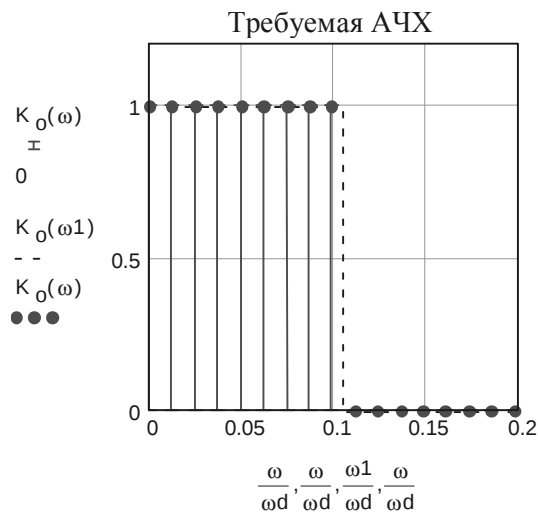
$$A1(\omega) := |H1(\exp(j \cdot \omega \cdot T))|$$

$$A2(\omega) := |H2(\exp(j \cdot \omega \cdot T))|$$

$$\text{step} := \frac{\omega d}{N}$$

$$\omega := 0, \text{step} \dots .2 \cdot \omega d$$

$$\omega 1 := 0, 0.0001 \cdot \omega d \dots .2 \cdot \omega d$$



3. ЦИФРОВАЯ ОБРАБОТКА ИЗОБРАЖЕНИЙ

В нашем мире существует множество сигналов, которые по своей природе являются двумерными и многомерными. Это, прежде всего различные изображения, получаемые при аэрофотосъемках и метеосъемках, рентгеновские снимки и сейсмограммы, фототелеграфные и телевизионные изображения, радиолокационные изображения, астрофизические и океанографические карты. Обработкой таких сигналов и занимается сравнительно новое направление в науке, называемое цифровой обработкой изображений и многомерных сигналов.

Без преувеличения можно сказать, что обработкой изображений в той или иной мере занимаются специалисты практически всех областей знаний. Трудно назвать такую область науки и техники, где бы ни использовалась обработка изображений. Это, прежде всего, связано с технической революцией в области обработки изображений, обусловленной с бурным развитием вычислительной техники.

Цифровая обработка изображений (ЦОИ) впервые была применена в связи с восстановлением и улучшением качества фотоснимков обратной стороны Луны и поверхности Марса.

В настоящее время с помощью ЦОИ решены такие сложные математические задачи, как идентификация личности по отпечаткам пальцев, по фотографии, восстановление изображений по проекциям в томографии (используется в медицинской диагностике для получения изображений внутренних органов) и др.

С общей точки зрения, казалось бы, не существует принципиальной разницы между цифровой обработкой одномерных и многомерных сигналов (изображений и полей).

В любом случае обработка включает в себя дискретизацию, квантование, выполнение различных ортогональных преобразований и сверток. Однако более глубокий анализ показывает, что переход от одномерного случая к двумерному является не только количественным, но и качественным. Многие проблемы, с которыми приходится сталкиваться при обработке многомерных сигналов, просто не существуют в обработке одномерных сигналов и, наоборот, многие трудности цифровой обработки сигналов отсутствуют или легко снимаются при переходе от одномерных сигналов к многомерным. Поэтому изучение основ цифровой обработки многомерных сигналов полезно не только специалистам, имеющим дело с обработкой изображений, но и специалистам, занимающимся обработкой одномерных сигналов в радиоэлектронике и системах управления.

На рис. 3.1 приведена типичная организация системы ЦОИ.

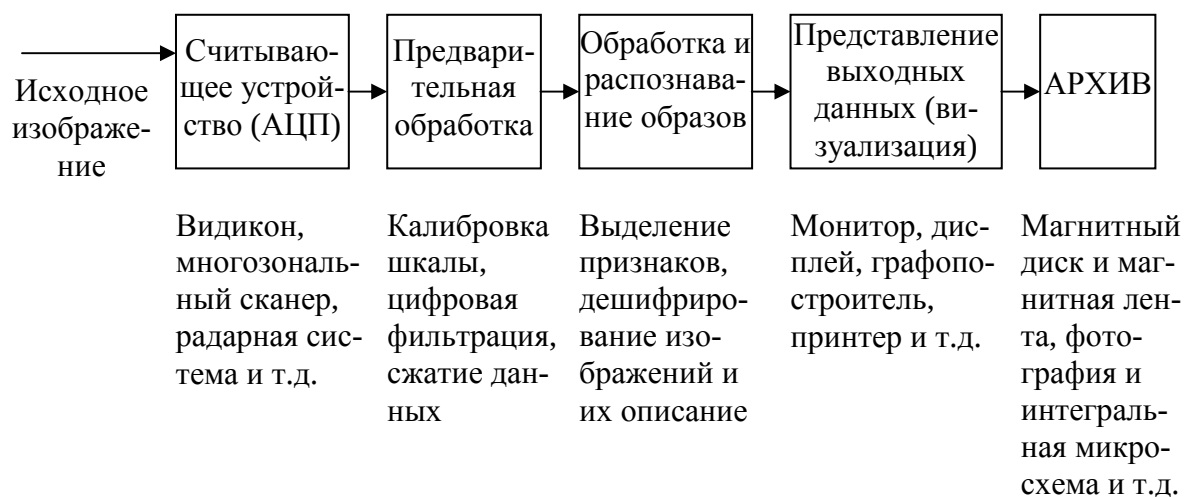


Рис. 3.1. – Типичная организация системы ЦОИ

При описании изображений вместо $x(nT)$ будем пользоваться обозначением двумерного сигнала $x(n_1, n_2)$ (рис. 3.2).

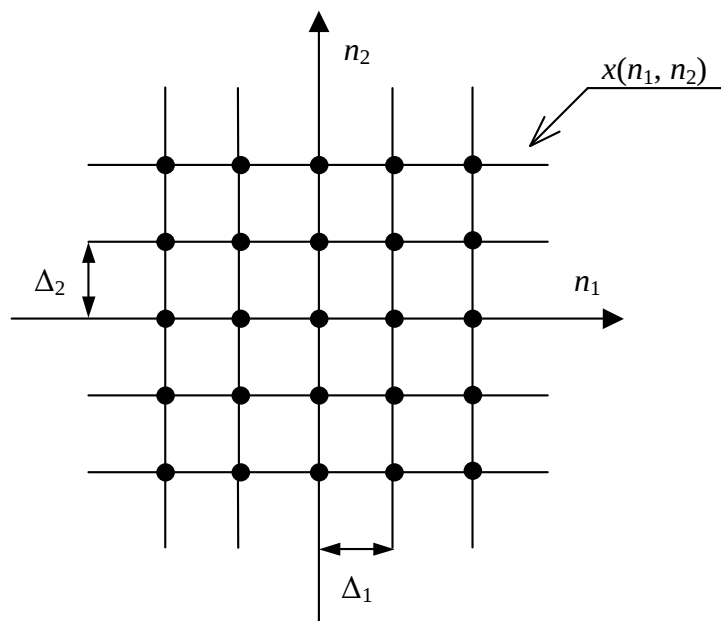


Рис. 3.2. – Сетка дискретизации двумерного сигнала

Дискретизация изображения:

Δ_1 – интервал дискретизации по n_1 ;

Δ_2 – интервал дискретизации по n_2 ;

Если $\Delta_1 = \Delta_2 = \Delta$, то сетка дискретизации имеет квадратную форму;

если $\Delta_1 \neq \Delta_2$, но $\Delta_1 = \text{const}$, $\Delta_2 = \text{const}$ – прямоугольная сетка дискретизации;

если Δ_1, Δ_2 – переменные величины, то случайная или квазислучайная сетка.

Кроме этого возможна треугольная и гексагональная дискретизация исходного изо-

бражения (с постоянным шагом).

Изображения бывают:

- истинные (оптические);
- синтезированные – радиолокационные (выборка по азимуту и углу), сейсмические и т.д.;
- «условные», т.е. любой двумерный массив данных (память ЭВМ).

К последним применяют, как правило, нелинейные алгоритмы сортировки, перемещений, сжатия информации или децимации и т.д.

Изображения можно разделить на статические:

- фотография;
- изображение, снятое сканером;
- отдельные кадры телевизионного изображения;

и динамические:

- кинофильмы;
- последовательности кадров ТВ изображения.

Статические изображения имеют две степени свободы (n_1, n_2), а динамические – три степени свободы: n_1, n_2 – определяют пространственные координаты, а n_3 – время (рис. 3.3).

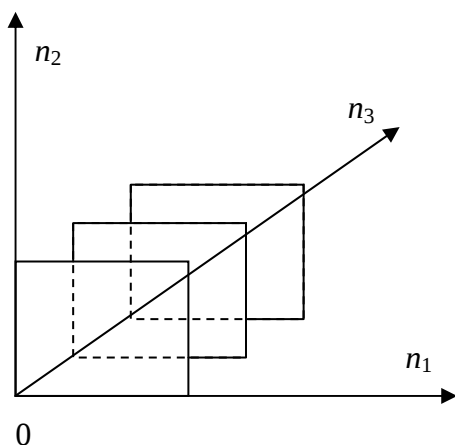


Рис. 3.3.

В связи с этим различают:

- внутрикадровые алгоритмы обработки и анализа изображений (используются для статических изображений);
- межкадровые алгоритмы обработки и анализа (используются для динамических изображений).

Внутрикадровые алгоритмы относят к пространственной обработке изображений, а межкадровые – к временной обработке.

Наиболее ярким примером межкадровой обработки является межкадровая разность, применяемая для обнаружения подвижных объектов.

вижных объектов.

Применяется также и пространственно-временная обработка изображений. Например, при обнаружении подвижных объектов используется предварительное сжатие изображения $512 \times 512 \rightarrow 64 \times 64$, а затем по сжатому изображению обнаруживают подвижные объекты.

Отличия ЦОИ от ЦОС:

1. При решении двумерных задач используются гораздо большие объемы данных, чем при решении одномерных задач (требуется высокое быстродействие при обработке сигналов в реальном времени). Так, например, в черно-белых телевизорах в одном кадре содержится порядка 500 тысяч элементов, поэтому, если для передачи 1 элемента использовать 1 байт (256 уровней), то объем переданной информации за одну секунду составит $500000 \times 25 = 12,5$ Мбайт. Такое количество информации необходимо обрабатывать за доли секунды. Отсюда возникает сложность аппаратной и программной реализации устройств для обработки изображений, поэтому процессы обработки распа-

раллеливают и обрабатывают изображения с помощью систолических процессоров.

2. Математическое описание многомерных систем недостаточно разработано из-за отсутствия адекватных моделей изображений. Поэтому в обработке изображений большую роль играют эвристические алгоритмы, которые зачастую дают лучшие результаты, чем оптимальные линейные алгоритмы, хорошо работающие при нормальном законе распределения помех и шумов.

3. Многомерные сигналы обладают большим числом степеней свободы, что дает большую гибкость при выборе алгоритмов обработки, а именно для буферизированных изображений, но в результате этого реализация и программирование более трудоемко.

Направления ЦОИ:

- улучшение визуального качества изображений (фильтрация помех и шумов);
- восстановление «смазанных» и расфокусированных изображений;
- спектральный анализ;
- оценка параметров изображений (координат, площадей, периметров), выделение признаков, классификация и распознавание изображений.

3.1. Представление и преобразование двумерных сигналов

На рис. 3.4 представлен двумерный бинарный сигнал ($\bullet$ – нуль, $\times$ – еди-

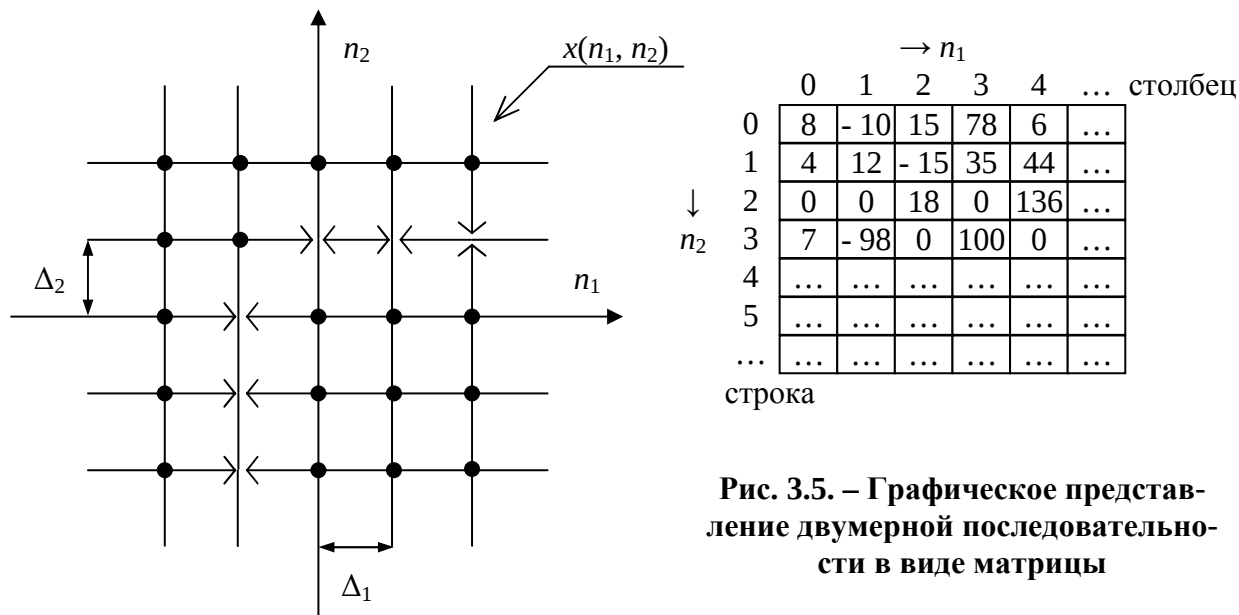


Рис. 3.4. – Графическое представление двумерных последовательностей ($\Delta_1 = \Delta_2 = \Delta = \text{const}$)

Рис. 3.5. – Графическое представление двумерной последовательности в виде матрицы

ница). Такое представление используется для изображения контуров, цифр и букв. Также сигнал можно представить в виде матрицы чисел (такое представление

используется, например, в телевидении), как показано на рис. 3.5. При графическом представлении двумерной последовательности отсчеты могут принимать любые значения, а не только бинарные.

Подобные представления двумерных сигналов необходимы для выбора алгоритма обработки изображения или осмысления результата обработки.

3.1.1. Некоторые особые двумерные последовательности

Двумерный единичный импульс (единичный отсчет) – ДЕИ:

$$\delta(n_1, n_2) = \begin{cases} 1, & n_1 = n_2 = 0, \\ 0 & \text{иначе.} \end{cases}$$

Двумерный линейный импульс (ДЛИ):

$$\delta_x(n_1) = \begin{cases} 1, & n_1 = 0, \\ 0 & \text{иначе.} \end{cases}$$

$$\delta_y(n_2) = \begin{cases} 1, & n_2 = 0, \\ 0 & \text{иначе.} \end{cases}$$

$$\delta(n_1, n_2) = \delta_x(n_1) \cdot \delta_y(n_2).$$

Двумерная единичная ступенька (ДЕС):

$$U(n_1, n_2) = \begin{cases} 1, & n_1 \geq 0, n_2 \geq 0, \\ 0 & \text{иначе.} \end{cases}$$

Двумерные экспоненциальные последовательности:

$$x(n_1, n_2) = a^{n_1} \cdot b^{n_2}, \quad -\infty < n_1 < \infty, \quad -\infty < n_2 < \infty,$$

если $a = \exp\{j\omega_1\}$ и $b = \exp\{j\omega_2\}$, то:

$$x(n_1, n_2) = \exp(j\omega_1 n_1 + j\omega_2 n_2) = \cos(\omega_1 n_1 + \omega_2 n_2) + j \sin(\omega_1 n_1 + \omega_2 n_2),$$

где ω_1 – пространственная частота по координате n_1 , а ω_2 – по координате n_2 . В этом случае экспоненциальная последовательность становится комплексной гармонической последовательностью. При $a = b = 1/2$ вид экспоненциальной последовательности приведен на рис. 3.6:

Последовательности называются разделимыми (сепарабельными), если они представлены в виде произведения функций каждого из переменных:

$$x(n_1, n_2) = x(n_1) \cdot x(n_2).$$

$\rightarrow n_1$			
1	0,5	0,25	0,125

$\downarrow n_2$

↓	1
	0,5
	0,25
	0,125

→

$\downarrow n_2$

	$\rightarrow n_1$			
	0	1	2	3
0	1	0,5	0,25	0,125
1	0,5	0,25	0,125	0,062
2	0,25	0,125	0,062	0,031
3	0,125	0,062	0,031	0,015

Рис. 3.6. – Двумерная экспоненциальная последовательность

В явном виде таких последовательностей существует не столь много. Однако любое двумерное множество с конечным числом ненулевых отсчетов можно записать в виде суммы ко-

нечного числа сепарабельных последовательностей

$$x(n_1, n_2) = \sum_{i=1}^N x_{i1}(n_1) \cdot x_{i2}(n_2),$$

где N – число ненулевых строк или столбцов.

Простейшее представление такого рода можно получить, выразив изображение в виде суммы отдельных строк последовательности. Для этого следует принять $x_{i1}(n_1) = x(n_1, i)$ и $x_{i2}(n_2) = \delta_y(n_2 - i)$.

Сепарабельные последовательности часто используются в качестве тестовых сигналов при оценке характеристик и при настройке экспериментальных систем.

Другим важным классом дискретных сигналов являются двумерные последовательности конечной протяженности. Слова «конечная протяженность» означают, что сигналы равны нулю вне области конечной протяженности в (n_1, n_2) – плоскости. Наиболее ярким примером последовательности

конечной протяженности является последовательность, получившаяся в результате пространственного стробирования (рис. 3.7).

Здесь мы имеем сигнал, отличный от нуля внутри области

$$n_1' \leq n_1 \leq n_1'', \quad n_2' \leq n_2 \leq n_2'',$$

$$N_1 = n_1'' - n_1', \quad N_2 = n_2'' - n_2',$$

где N_1 и N_2 – размеры опорной области.

Область, внутри которой значения сигнала отличны от нуля, называется опорной областью сигнала.

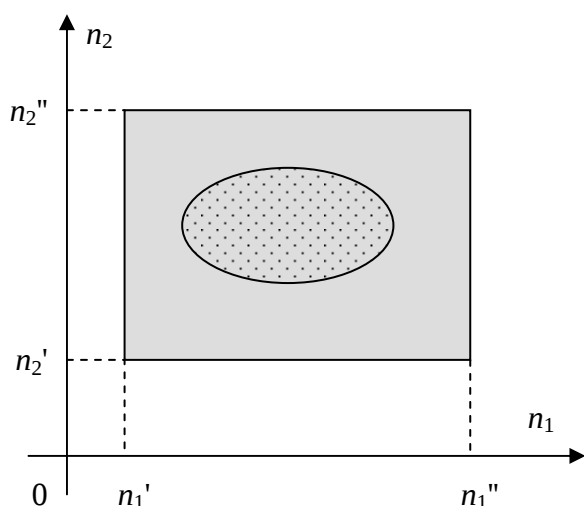


Рис. 3.7.

3.1.2. Многомерные системы

Для обработки и преобразования подобных сигналов служат многомерные системы. Формально многомерная система это оператор, отображающий один (входной) сигнал на другой (выходной) сигнал:

$$y(n_1, n_2) = T[x(n_1, n_2)].$$

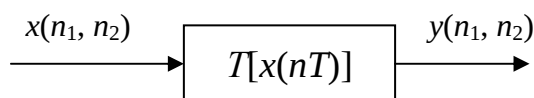


Рис. 3.8. – Многомерная система

Оператор $T[\cdot]$ представляет собой правило или набор правил, по которым происходит отображение входного сигнала на выходной, или даже таблицу соответствия выходных сигналов различным входным.

3.1.3. Базовые операции, используемые в многомерных системах

Рассмотрим несколько основных операций, которые выступают в качестве «кирпичиков» при разработке более сложных систем:

1. Сложение (вычитание) изображений

$$z(n_1, n_2) = x(n_1, n_2) \pm y(n_1, n_2), \text{ где } x, y, z - \text{дискретные сигналы.}$$

2. Умножение изображения на константу (нормирование)

$$y(n_1, n_2) = c \cdot x(n_1, n_2).$$

3. Сдвиг изображения (сдвиг двумерных последовательностей)

$$y(n_1, n_2) = x(n_1 - m_1, n_2 - m_2), \text{ где } (m_1, m_2) - \text{величина сдвига.}$$

4. Пространственное стробирование (маскирование сигнала)

$$y(n_1, n_2) = c(n_1, n_2) \cdot x(n_1, n_2).$$

Совокупность чисел $c(n_1, n_2)$ описывает маску или строб. Эту совокупность чисел можно также рассматривать как двумерную последовательность. В этом случае правая часть равенства представляет собой поэлементное произведение двух последовательностей.

5. Безынерционные нелинейные преобразования (имеется в виду «попиксельное» преобразование каждого отсчета двумерной последовательности).

$$\text{Возведение в квадрат: } y(n_1, n_2) = x^2(n_1, n_2);$$

$$\text{Взятие модуля: } y(n_1, n_2) = |x(n_1, n_2)|.$$

3.1.4. Линейные и инвариантные к сдвигу многомерные системы

Наиболее важным классом многомерных систем является класс линейных и инвариантных к сдвигу систем.

Систему называют линейной, если выполняется условие:

пусть

$$y_1(n_1, n_2) = L[x_1(n_1, n_2)], \quad y_2(n_1, n_2) = L[x_2(n_1, n_2)],$$

тогда

$$L[a \cdot x_1(n_1, n_2) + b \cdot x_2(n_1, n_2)] = a \cdot y_1(n_1, n_2) + b \cdot y_2(n_1, n_2),$$

где a и b – константы; x_1, x_2, y_1, y_2 – входные и выходные последовательности, соответственно.

Линейные системы подчиняются принципу суперпозиции. Отклик линейной системы на взвешенную сумму входных сигналов равен взвешенной сумме откликов на отдельные входные сигналы. Для системы, инвариантной к сдвигу, должно выполняться следующее условие:

пусть $y(n_1, n_2) = T[x(n_1, n_2)]$, тогда для инвариантной к сдвигу системы выполняется условие $T[x(n_1 - m_1, n_2 - m_2)] = y(n_1 - m_1, n_2 - m_2)$ для всех последовательностей x и для всех целочисленных сдвигов (m_1, m_2) .

Линейность и инвариантность к сдвигу являются независимыми свойствами системы. Ни одно из этих свойств не подразумевает обязательного наличия другого.

$L[x(n_1, n_2)] = c(n_1, n_2) \cdot x(n_1, n_2)$ – пространственное маскирование линейно, но не инвариантно к сдвигу.

$T[x(n_1, n_2)] = [x(n_1, n_2)]^2$ – эта система инвариантна к сдвигу, но нелинейная.

Линейные инвариантные к сдвигу дискретные системы (ЛИС-системы) – это наиболее часто используемый класс систем для обработки дискретных сигналов любой размерности. Эти системы отличаются простотой, как при разработке, так и при анализе, но в тоже время они обладают достаточными возможностями для решения многих задач.

Выходной сигнал двумерной линейной системы равен:

$$y(n_1, n_2) = \sum_{i_1=0}^{\infty} \sum_{i_2=0}^{\infty} x(i_1, i_2) \cdot h_{i_1, i_2}(n_1, n_2), \text{ где } h_{i_1, i_2}(n_1, n_2) \text{ – отклик на ДЕИ в}$$

точке с координатами i_1 и i_2 .

Для систем, инвариантных к сдвигу (систем с постоянными параметрами): $y(n_1, n_2) = \sum_{i_1=0}^{\infty} \sum_{i_2=0}^{\infty} x(i_1, i_2) \cdot h(n_1 - i_1, n_2 - i_2)$.

Пример трехмерной системы. Пусть $x(n_1, n_2, n_3)$ – трехмерный сигнал, где n_1, n_2 – пространственные координаты, n_3 – номер кадра (см. рис. 3.3).

Простейшая обработка – межкадровая разность текущего и предыдущего кадров. Следовательно, $y(n_1, n_2, n_3) = x(n_1, n_2, n_3) - x(n_1, n_2, n_3 - 1)$. Данная обработка используется для обнаружения подвижных объектов.

3.2. Двумерные линейные фильтры

Схематически алгоритм двумерной линейной фильтрации приведен на рис. 3.9.

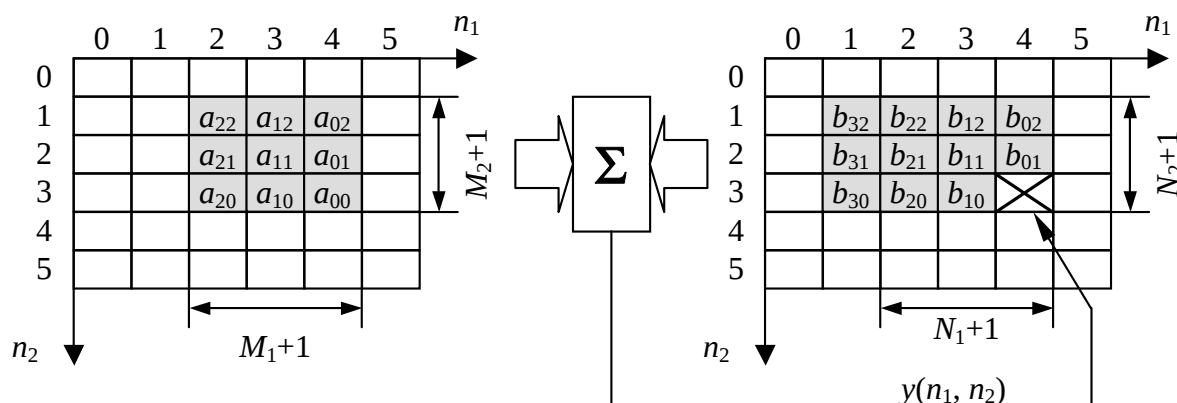


Рис. 3.9. – Алгоритм двумерной линейной фильтрации

Здесь $(M_1+1)(M_2+1)$ – размер опорной области по входным данным – $x(n_1, n_2)$, а $(N_1+1)(N_2+1)$ – размер опорной области по выходным данным – $y(n_1, n_2)$.

Разностное уравнение:

$$y(n_1, n_2) = \sum_{i_1=0}^{M_1} \sum_{i_2=0}^{M_2} a_{i_1, i_2} x(n_1 - i_1, n_2 - i_2) + \sum_{j_1=0}^{N_1} \sum_{j_2=0}^{N_2} b_{j_1, j_2} y(n_1 - j_1, n_2 - j_2),$$

где $n_1 \geq 0, n_2 \geq 0, (j_1, j_2) \neq 0$.

Пример сглаживающего фильтра:

$$y(n_1, n_2) = x(n_1, n_2) + k y(n_1 - 1, n_2 - 1), \quad k = b_{11}, \quad n_1 \geq 0, \quad n_2 \geq 0.$$

Импульсная характеристика:

$$h(n_1, n_2) = y(n_1, n_2) |_{x(n_1, n_2) = \delta(n_1, n_2)} \text{ – при нулевых начальных условиях.}$$

Системная функция:

$$H(z_1, z_2) = \frac{Y(z_1, z_2)}{X(z_1, z_2)} \text{ – системная функция определяется как отношение } z\text{-образов входной и выходной последовательностей; здесь } z_1^{-1} \text{ – задержка на 1 шаг по строке (координата } n_1); z_2^{-1} \text{ – задержка на 1 шаг по кадру или столбцу (координата } n_2).$$

$F(z_1, z_2) = \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} f(n_1, n_2) z_1^{-n_1} z_2^{-n_2}$ – двумерное Z-преобразование.

$$H(z_1, z_2) = \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} h(n_1, n_2) z_1^{-n_1} z_2^{-n_2} \text{ – системная функция в Z-форме.}$$

Если заданы коэффициенты a_{i_1, i_2} и b_{j_1, j_2} в разностном уравнении (РУ), то системная функция (СФ) примет вид:

$$H(z_1, z_2) = \frac{\sum_{i_1=0}^{M_1} \sum_{i_2=0}^{M_2} a_{i_1, i_2} z_1^{-i_1} z_2^{-i_2}}{1 - \sum_{j_1=0}^{N_1} \sum_{j_2=0}^{N_2} b_{j_1, j_2} z_1^{-j_1} z_2^{-j_2}}; \text{ отсчет } (j_1, j_2) = 0 \text{ исключается.}$$

Очень широкое применение в обработке изображений находят нерекурсивные фильтры (НРФ):

$$\text{РУ: } y(n_1, n_2) = \sum_{i_1=0}^{M_1} \sum_{i_2=0}^{M_2} a_{i_1, i_2} x(n_1 - i_1, n_2 - i_2);$$

$$\text{СФ: } H(z_1, z_2) = \sum_{i_1=0}^{M_1} \sum_{i_2=0}^{M_2} a_{i_1, i_2} z_1^{-i_1} z_2^{-i_2}.$$

Выходной сигнал двумерного нерекурсивного фильтра представляется в виде линейной комбинации текущего и предыдущих отсчетов входного сигнала. Структурная схема подобного фильтра показана на рисунке 3.10.

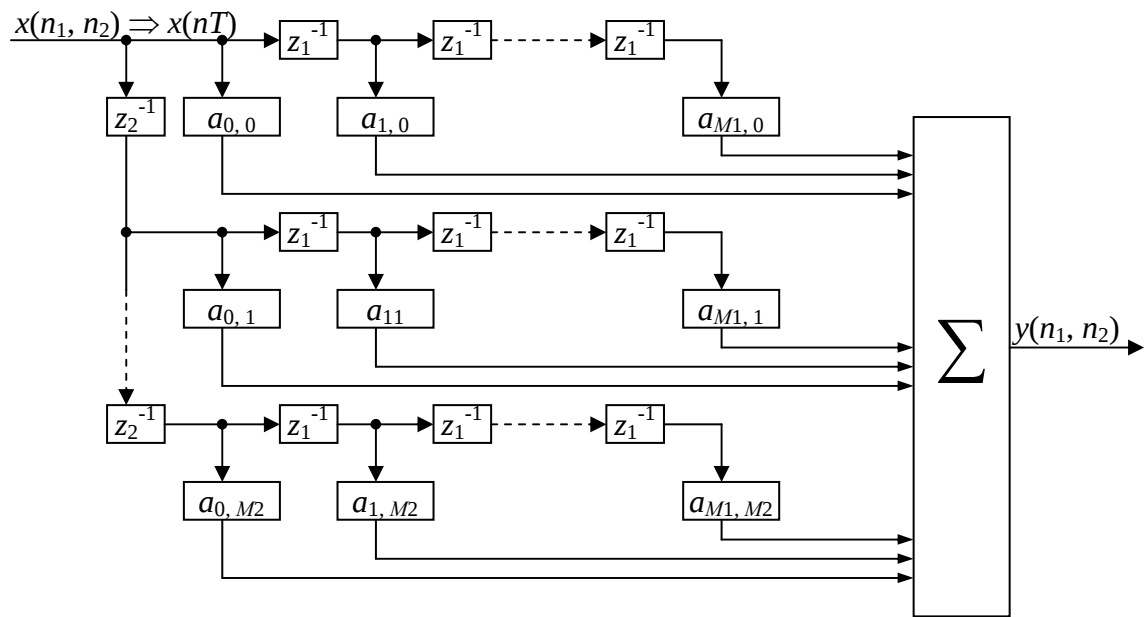


Рис. 3.10. – Структурная схема нерекурсивного фильтра

Коэффициенты фильтра задаются в виде двумерных масок (рис. 3.11).

Рассмотрим типы наиболее используемых масок

на примере маски размером 3x3 (рис. 3.12):

а) скользящее среднее;

б) лапласиан;

в) двойное дифференцирование;

г) оператор выделения вертикальных линий;

д) оператор выделения малоразмерных деталей

из шумов;

е) градиентный оператор выделения перепада.

a_{22}	a_{12}	a_{02}
a_{21}	a_{11}	a_{01}
a_{20}	a_{10}	a_{00}

Рис. 3.11. – Маска (3x3)

а)

1	1	1
1	1	1
1	1	1

б)

-1	-1	-1
-1	8	-1
-1	-1	-1

в)

1	-2	1
-2	4	-2
1	-2	1

г)

-1	2	-1
-1	2	-1
-1	2	-1

д)

1	2	1
2	4	2
1	2	1

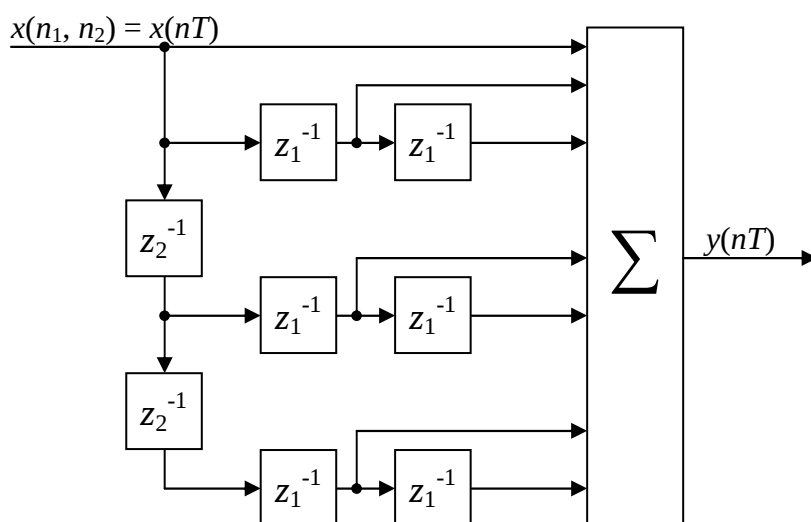
е)

1	1	-1
1	-2	-1
1	1	-1

Рис. 3.12. – Типы масок

3.2.1 Оператор двумерного «скользящего среднего»

Структурная схема фильтра «скользящее среднее» приведена на рис. 3.13. Часто для нормирования делят сумму отсчетов на сумму коэффициентов, т.е. на 9. Коэффициенты $a_{i_1, i_2} = 1$ (рис. 3.14).



1	1	1
1	1	1
1	1	1

Рис. 3.14.

Если входной массив буферизирован, то сам генератор адреса реализует задержки.

Рис. 3.13. – Структурная схема фильтра «скользящее среднее»

Разностное уравнение:

$$y(n_1, n_2) = x(n_1, n_2) + x(n_1 - 1, n_2) + x(n_1 - 2, n_2) + x(n_1, n_2 - 1) + x(n_1 - 1, n_2 - 1) + \\ + x(n_1 - 2, n_2 - 1) + x(n_1, n_2 - 2) + x(n_1 - 1, n_2 - 2) + x(n_1 - 2, n_2 - 2).$$

Импульсная характеристика такого фильтра $h(n_1, n_2) = h(n_1) * h(n_2)$ – сепарабельна (разделима), т. к. маска симметрична (рис. 3.15).

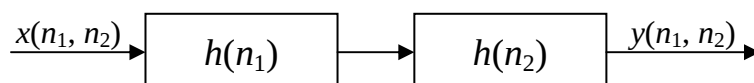


Рис. 3.15.

3.2.2. Лапласиан для «восьми соседей»

На рис. 3.16 показан дифференцирующий оператор, реализованный на основе маски (3x3).

Незаполненные маски имеют дифференцирование не такое «жесткое». Их применяют для выделения малоразмерных объектов, расположенных на гладких фонах. В них размер маски согласован с размером объекта, а именно: размер маски выбирается в 2 раза большим, чем размер объекта. Пример незаполненной маски 3x3 показан на рис. 3.17.

-1	-1	-1
-1	8	-1
-1	-1	-1

Рис. 3.16.

0	-1	0
-1	4	-1
0	-1	0

Рис. 3.17.

Лапласианы получают путем сложения дифференцирующих масок.

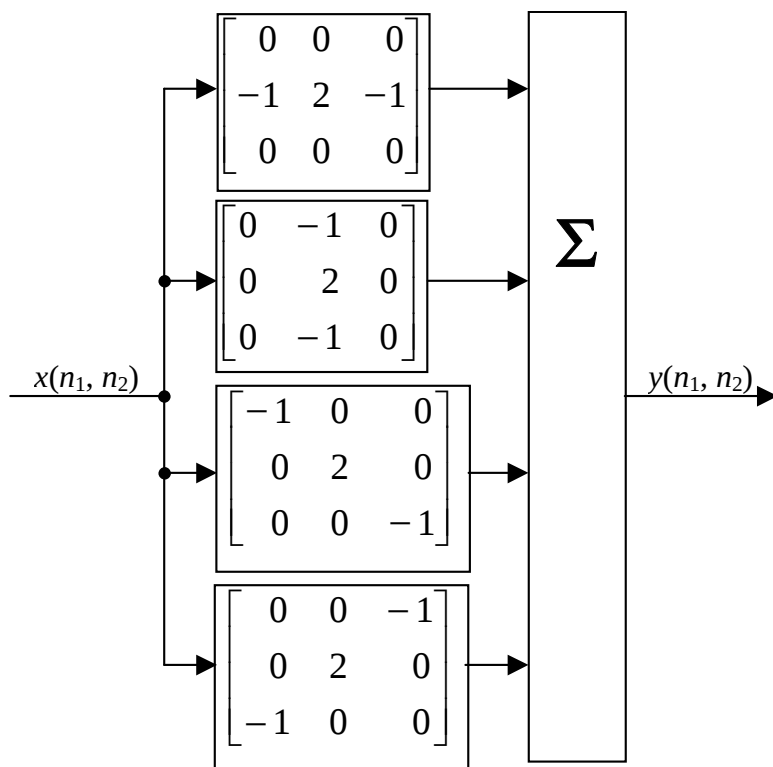


Рис. 3.18. – Структурная схема лапласиана для «восьми соседей»

Сложив первые два блока, получим лапласиан для «четырёх соседей», а сложив все четыре маски, получим лапласиан для восьми соседей (рис. 3.18).

Добавив к исходному сигналу сигнал после обработки лапласианом, получим обработку, аналогичную двумерному апертурному корректору, применяемому в телевидении для повышения резкости изображения.

Реализация лапласиана для «восьми соседей» (рис. 3.19). Центральный элемент берется с коэффициентом 9.

Разностное уравнение лапласиана для «четырёх со-

седей» (рис. 3.20) имеет вид

$$y(n_1, n_2) = -x(n_1 - 1, n_2) - x(n_1, n_2 - 1) + 4x(n_1 - 1, n_2 - 1) - x(n_1 - 2, n_2 - 1) - x(n_1 - 1, n_2 - 2).$$

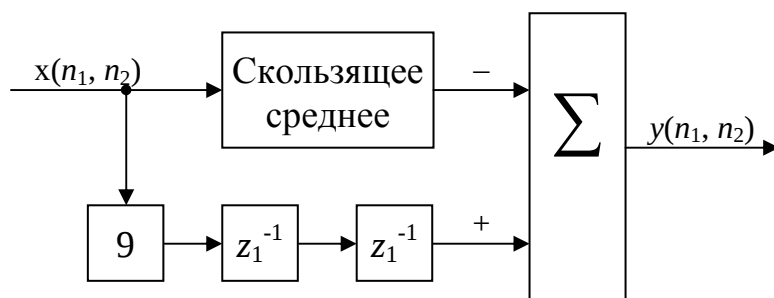


Рис. 3.19.

0	-1	0
-1	4	-1
0	-1	0

Рис. 3.20.

3.2.3. Оператор выделения линий (контуров) в изображении

Виды масок для операторов выделения контуров в изображении приведены на рис. 3.21 – 3.22:

а) ВГ – оператор выделения вертикальной границы.

Вторая производная в направлении n_1 . Вертикальная граница проходит с коэффициентом усиления $K = 6$.

$$\begin{array}{|c|c|c|} \hline -1 & 2 & -1 \\ \hline -1 & 2 & -1 \\ \hline -1 & 2 & -1 \\ \hline \end{array} = \begin{array}{|c|c|c|} \hline -1 & 2 & -1 \\ \hline \end{array} ** \begin{array}{|c|} \hline 1 \\ \hline 1 \\ \hline 1 \\ \hline \end{array}$$

Рис. 3.21

б) ГГ – оператор выделения горизонтальной границы.

$$\begin{array}{|c|c|c|} \hline -1 & -1 & -1 \\ \hline 2 & 2 & 2 \\ \hline -1 & -1 & -1 \\ \hline \end{array} = \begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline \end{array} ** \begin{array}{|c|} \hline -1 \\ \hline 2 \\ \hline -1 \\ \hline \end{array}$$

Рис. 3.22

в) Выделение правой диагонали (оператор неразделим) – рис. 3.23.

$$\begin{array}{|c|c|c|} \hline -1 & -1 & 2 \\ \hline -1 & 2 & -1 \\ \hline 2 & -1 & -1 \\ \hline \end{array}$$

Рис. 3.23

г) Оператор (маска) выделения левой диагонали (оператор неразделим) – рис. 3.24.

$$\begin{array}{|c|c|c|} \hline 2 & -1 & -1 \\ \hline -1 & 2 & -1 \\ \hline -1 & -1 & 2 \\ \hline \end{array}$$

Рис. 3.24

3.2.4. Оператор двойного дифференцирования

Оператор двойного дифференцирования используется для выделения малоразмерных деталей в изображении (рис. 3.25).

$$\begin{array}{|c|c|c|} \hline -1 & -2 & -1 \\ \hline -2 & 4 & -2 \\ \hline -1 & -2 & -1 \\ \hline \end{array} = \begin{array}{|c|c|c|} \hline -1 & 2 & -1 \\ \hline \end{array} ** \begin{array}{|c|} \hline -1 \\ \hline 2 \\ \hline -1 \\ \hline \end{array}$$

Рис. 3.25

Пример незаполненной маски 5x5 для двойного дифференцирования (рис. 3.26):

1	0	-2	0	1
0	0	0	0	0
-2	0	4	0	-2
0	0	0	0	0
1	0	-2	0	1

Рис. 3.26

3.2.5. Оператор выделения малоразмерных объектов из шумов и фонов, подчиняющихся нормальному закону распределения

Маска интегрирующего типа имеет вид, приведенный на рис. 3.27.

$$\begin{array}{|c|c|c|} \hline 1 & 2 & 1 \\ \hline 2 & 4 & 2 \\ \hline 1 & 2 & 1 \\ \hline \end{array} = \begin{array}{|c|c|c|} \hline 1 & 2 & 1 \\ \hline \end{array} ** \begin{array}{|c|} \hline 1 \\ \hline 2 \\ \hline 1 \\ \hline \end{array}$$

Рис. 3.27

Физический смысл – выделение малоразмерных объектов на фоне «белого» шума.

Вид маски согласован с видом сигнала, т.е. маска есть образ сигнала.

3.2.6. Градиентные операторы выделения перепадов яркости в различных пространственных направлениях

В лапласианах используются вторые частные производные, а в градиентных операторах – первые частные производные в различных пространственных направлениях (рис. 3.28 – 3.30).

Оператор «Запад»

1	1	-1
1	-2	-1
1	1	-1

Рис. 3.28

Оператор «Север»

1	1	1
1	-2	1
-1	-1	-1

Рис. 3.29

Оператор «Юго-запад»

1	-1	-1
1	-2	-1
1	1	1

Рис. 3.30

3.3. Рекурсивная обработка изображений

На рис. 3.31 показан процесс формирования выходного сигнала в двумерном рекурсивном фильтре. Сканирование слева направо и сверху вниз.

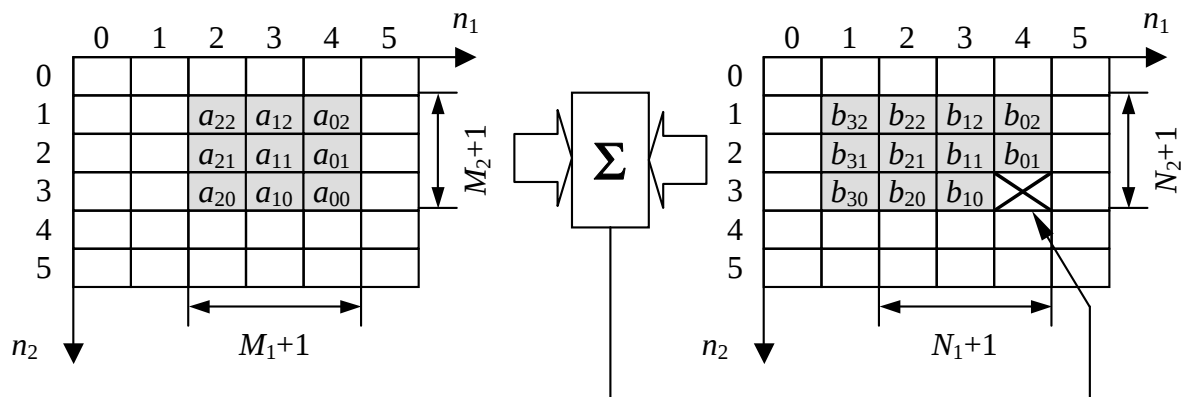


Рис. 3.31. – Процесс формирования выходного сигнала в двумерном рекурсивном фильтре (сканирование слева направо и сверху вниз)

Известные алгоритмы обработки телевизионных изображений, построенные на базе нерекурсивной СМА, характеризуются большими аппаратными затратами и низкой производительностью. Общий подход к повышению универсальности и производительности цифровых фильтров основан на распараллеливании операций обработки сигнала. В частности, импульсную характеристику двумерного фильтра, представленную конечным числом ненулевых отсчетов, можно записать в виде

$$h(n_1, n_2) = \sum_{i=1}^K h_{i1}(n_1) ** h_{i2}(n_2),$$

здесь «**» обозначает операцию двумерной дискретной свертки. В том случае, если импульсную характеристику фильтра можно представить в виде дискретной свертки функции горизонтальных и вертикальных индексов при $K = 1$, то такой фильтр называют разделимым (*сепарабельным*). Общее число операций в апертуре при вычислении отсчетов сигнала в разделимом фильтре равно $(M_1 + M_2)$, а в неразделимом фильтре общее число операций в апертуре значительно больше и достигает числа $M_1 \cdot M_2$. Если одномерные импульсные характеристики $h_{i1}(n_1)$ и $h_{i2}(n_2)$ реализуются с помощью рекурсивных процедур, то число операций при вычислении выходных отсчетов двумерного фильтра можно сделать много меньшим чем $(M_1 + M_2)$ и не зависящем от размера апертур. Такой подход к проектированию двумерных цифровых фильтров, особенно при больших размерах маски (M_1 и M_2), реализует существенный выигрыш по аппаратным затратам и повышает скорость вычислений.

При обработке изображений и оценке их информативных параметров часто используются процедуры с конечными импульсными характеристиками (КИХ), например, при оценке фонового сигнала в текущем фрагменте изображения, выделении мелких деталей в изображении их пространственного положения. Наиболее производительными и эффективными процедурами обработки телевизионных изображений в реальном времени являются алгоритмы, реализуемые в рекурсивных сепарабельных цифровых фильтрах (РСЦФ). Положительный эффект в РСЦФ достигается за счет наиболее полного согласования их работы с характером растрового съема телевизионной информации.

3.3.1. Интервальное интегрирование

Основной базовой ячейкой обработки видеосигнала можно выбрать рекурсивный фильтр, выполняющий операцию интегрирования на заданном интервале (рис. 3.32).

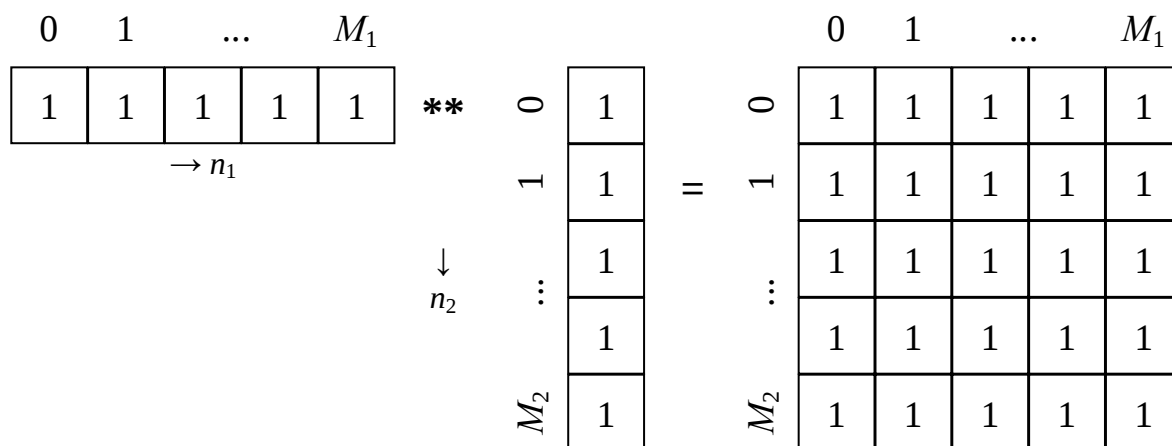


Рис. 3.32. – Рекурсивный алгоритм "скользящее среднее"

Работа фильтра описывается разностным уравнением

$$y_1(nT) = ax(nT) - bx(nT - MT) + cy_1(nT - T).$$

При различных значениях коэффициентов в этом выражении получаем различные формы и длительности КИХ в заданном направлении фильтрации. Включая каскадно такие рециркулирующие ячейки, можно получить достаточно широкий класс двумерных КИХ.

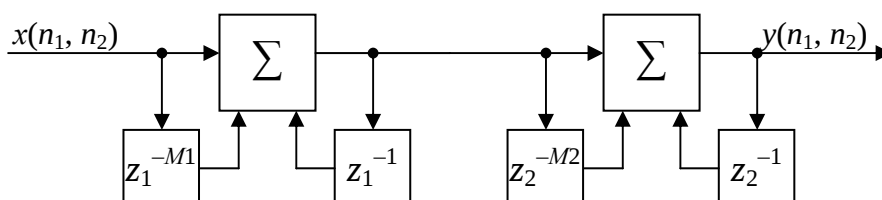


Рис. 3.33

Например (рис. 3.33), при последовательном включении двух рециркуляторов согласно приведенной выше формуле, обрабатывающих сигнал вдоль строки и кадра, при $a = b = c = 1$ получаем процедуру фильтрации, реализующую «скользящее среднее», то есть двумерное интегрирование, причем число арифметических операций равно четырем и не зависит от размера фильтрующей маски. Из рисунка видно, что структура фильтра существенно упростилась. Вместо сложных линий с отводами M_1 и M_2 используются более простые линии задержки. При этом структуры сумматоров существенно упрощаются. На рис. 3.33 M_1 – направление по элементам, M_2 – направление по строкам, $z_1^{-1} \rightarrow$ – задержка на один элемент T , z_2^{-1} – на одну строку T_c .

Используются две простые операции в первом каскаде и две во втором. Всего четыре операции. При малых M_1 и M_2 выигрыш по быстродействию незначительный, но при больших M_1 и M_2 выигрыш весьма существенен.

3.3.2. Интервальное дифференцирование

Данная обработка используется при дискриминировании малоразмерных объектов, когда производится вычисление сигнала рассогласования между центром строба и объекта (рис. 3.34).

Интервальное дифференцирование в заданном пространственном направлении реализуется в РСЦФ с линейной двухполярной КИХ, который описывается уравнением

$$y_2(nT) = \frac{1-M}{2} [x(nT) - x(nT - MT - T)] + \frac{1+M}{2} [x(nT - T) - x(nT - MT)] + 2y_2(nT - T) - y_2(nT - 2T).$$

$$\begin{array}{c}
 \begin{array}{ccccc}
 0 & 1 & \dots & M_1 \\
 \hline
 -2 & -1 & 0 & 1 & 2
 \end{array}
 \xrightarrow{n_1}
 \begin{array}{c}
 \text{**} \\
 \downarrow n_2 \\
 \begin{array}{c}
 0 \\
 1 \\
 \vdots \\
 M_2
 \end{array}
 \begin{array}{c}
 1 \\
 1 \\
 1 \\
 1 \\
 1
 \end{array}
 \end{array}
 =
 \begin{array}{c}
 \begin{array}{ccccc}
 0 & 1 & \dots & M_1 \\
 \hline
 -2 & -1 & 0 & 1 & 2
 \end{array}
 \begin{array}{c}
 0 \\
 1 \\
 \vdots \\
 M_2
 \end{array}
 \begin{array}{c}
 1 \\
 1 \\
 1 \\
 1 \\
 1
 \end{array}
 \end{array}
 \end{array}$$

Рис. 3.34.

Дифференцирующий РСЦФ подобного вида для формирования выходного результата выполняет всего лишь семь арифметических операций, при чем коэффициенты фильтра (с целью повышения быстродействия) выбираются кратными степени два. Фильтр в этом случае реализуется в виде двухкаскадного рециркулятора вида (п. 3.3.1) с дополнительными связями.

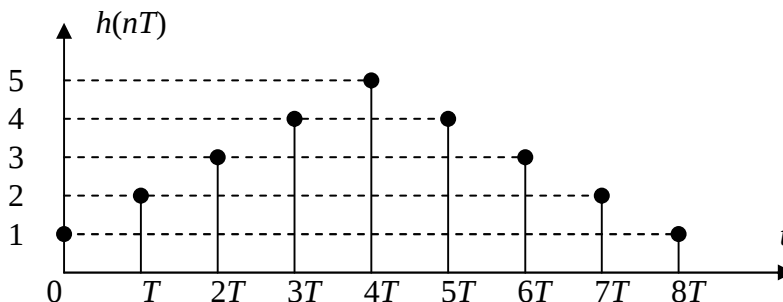
3.3.3. Квазиоптимальная фильтрация малоразмерных объектов из шумов

На рис. 3.35 показано прямоугольное окно (маска) 5x1. Если этим оператором обработать исходное изображение дважды, то получим ИХ треугольного вида (трижды – колоколообразного вида) (рис. 3.36).

1	1	1	1	1
---	---	---	---	---

Рис. 3.35

$$\begin{array}{|c|c|c|c|c|}
 \hline
 1 & 1 & 1 & 1 & 1 \\
 \hline
 \end{array}
 *
 \begin{array}{|c|c|c|c|c|}
 \hline
 1 & 1 & 1 & 1 & 1 \\
 \hline
 \end{array}
 =
 \begin{array}{|c|c|c|c|c|c|c|c|c|}
 \hline
 1 & 2 & 3 & 4 & 5 & 4 & 3 & 2 & 1 \\
 \hline
 \end{array}$$



$$\begin{array}{|c|c|c|c|c|}
 \hline
 1 & 1 & 1 & 1 & 1 \\
 \hline
 \end{array}
 *
 \begin{array}{|c|c|c|c|c|}
 \hline
 1 & 1 & 1 & 1 & 1 \\
 \hline
 \end{array}
 *
 \begin{array}{|c|c|c|c|c|}
 \hline
 1 & 1 & 1 & 1 & 1 \\
 \hline
 \end{array}
 =
 \begin{array}{|c|c|c|c|c|c|c|c|c|c|c|c|c|c|}
 \hline
 1 & 3 & 6 & 10 & 15 & 18 & 19 & 18 & 15 & 10 & 6 & 3 & 1 \\
 \hline
 \end{array}$$

Рис. 3.36

Следует отметить, что ИХ ЦФ после четырех или пяти проходов прямоугольной маской практически гауссоидальная.

Для квазиоптимальной фильтрации сигнала телевизионного изображения из шумов можно использовать РСЦФ, описываемый уравнением

$$y_3(nT) = x(nT) - 2x(nT - MT) + x(nT - 2MT) + 2y_3(nT - T) - y_3(nT - 2T).$$

При последовательном включении двух фильтров такого типа форма результирующей КИХ близка к форме сигнала от малой детали изображения, а такой фильтр близок к оптимальному.

3.3.4. Трапецеидальная ИХ

Для получения ИХ трапецеидального вида следует брать образующие прямоугольные окна различной длительности (рис. 3.37).

Пропустив еще раз через единичную маску 3×1 , получим колоколообразный спад. Этот алгоритм используется как при сглаживании, так и при моделировании (генерации) процессов. Следует отметить, что в природе наблюдается тенденция к «нормализации» процесса при многократном сглаживании.

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \end{bmatrix} * \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 & 3 & 3 & 2 & 1 \end{bmatrix}$$

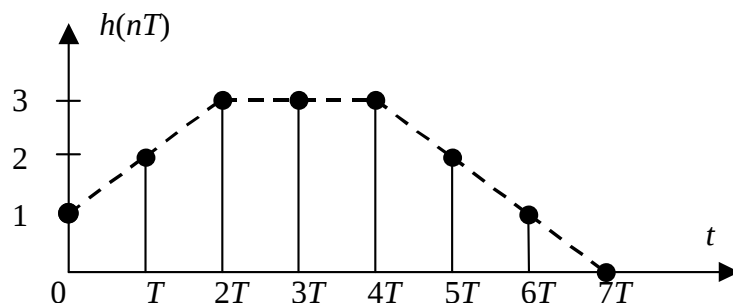


Рис. 3.37

3.3.5. Двухкаскадный РСЦФ

На рис. 3.38 приведена структурная схема двухкаскадного РСЦФ, который реализует алгоритмы интервального интегрирования, дифференцирования и квазиоптимальной обработки сигнала в одном из направлений фильтрации. Отметим, что на рис. 3.38 $y_1(nT)$ формируется постоянно, $y_2(nT)$ – при подаче на второй каскад рециркулятора выходного сигнала умножителя на M , $y_3(nT)$ – при подключении на вход второго рециркулятора выходного сигнала первого рециркулятора, задержанного на M тактов.

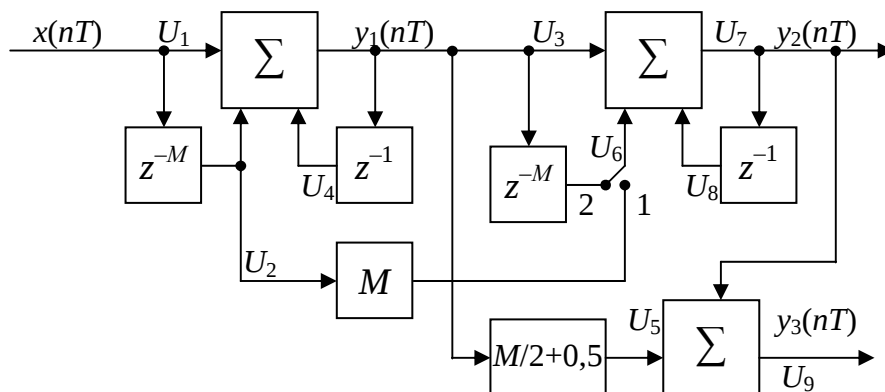


Рис. 3.38. – Структурная схема двухкаскадной рекурсивно-сепарабельной ячейки (обработка в одном из пространственных направлений)

Линейная комбинация описанных выше рециркуляторов при построении РСЦФ, а также задание необходимых направлений фильтрации позволяет на базе однотипных звеньев реализовать широкий класс двумерных КИХ. Следует отметить, что во всех предлагаемых РСЦФ организуется конвейерный принцип обработки видеоинформации, который характеризуется высоким быстродействием. Положительным моментом при использовании предлагаемых РСЦФ является их универсальность и малое число структурных компонент, необходимых для построения сложных двумерных фильтров.

На рис. 3.39 показаны локальные ИХ в каждой из точек приведенной выше схемы РСЦФ при различных положениях переключателя во втором каскаде фильтра.

Разностные уравнения:

1. Интервальное интегрирование (выход – U_3)

$$y_1(nT) = x(nT) - x(nT - MT) + y_1(nT - T).$$

2. Квазиоптимальная фильтрация (выход – U_7)

$$y_2(nT) = x(nT) - 2x(nT - MT) + x(nT - 2MT) + 2y_2(nT - T) - y_2(nT - 2T).$$

3. Интервальное дифференцирование (выход – U_9)

$$y_3(nT) = \frac{1-M}{2}[x(nT) - x(nT - MT - T)] + \frac{1+M}{2}[x(nT - T) - 2x(nT - MT)] + 2y_3(nT - T) + y_3(nT - 2T).$$

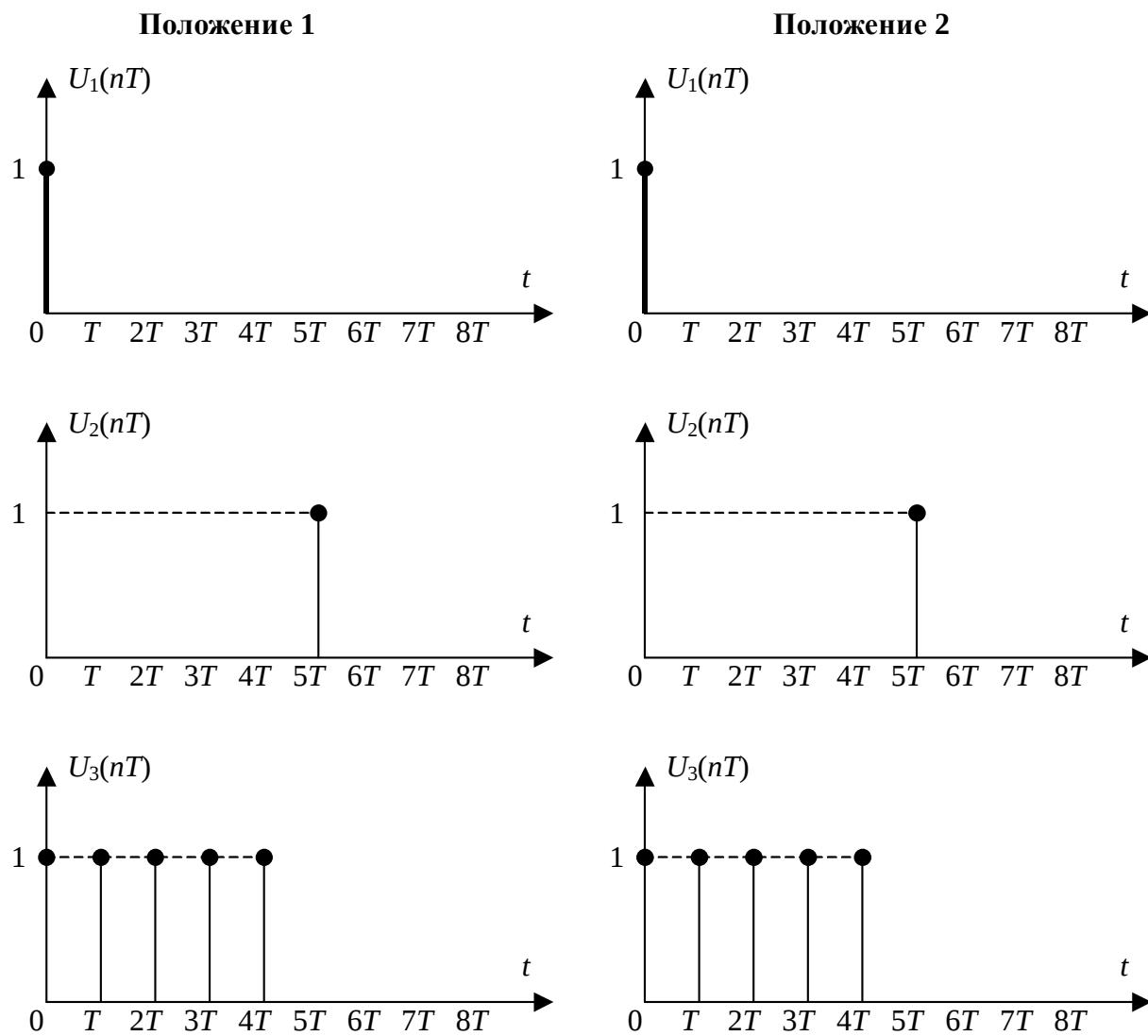


Рис. 3.39

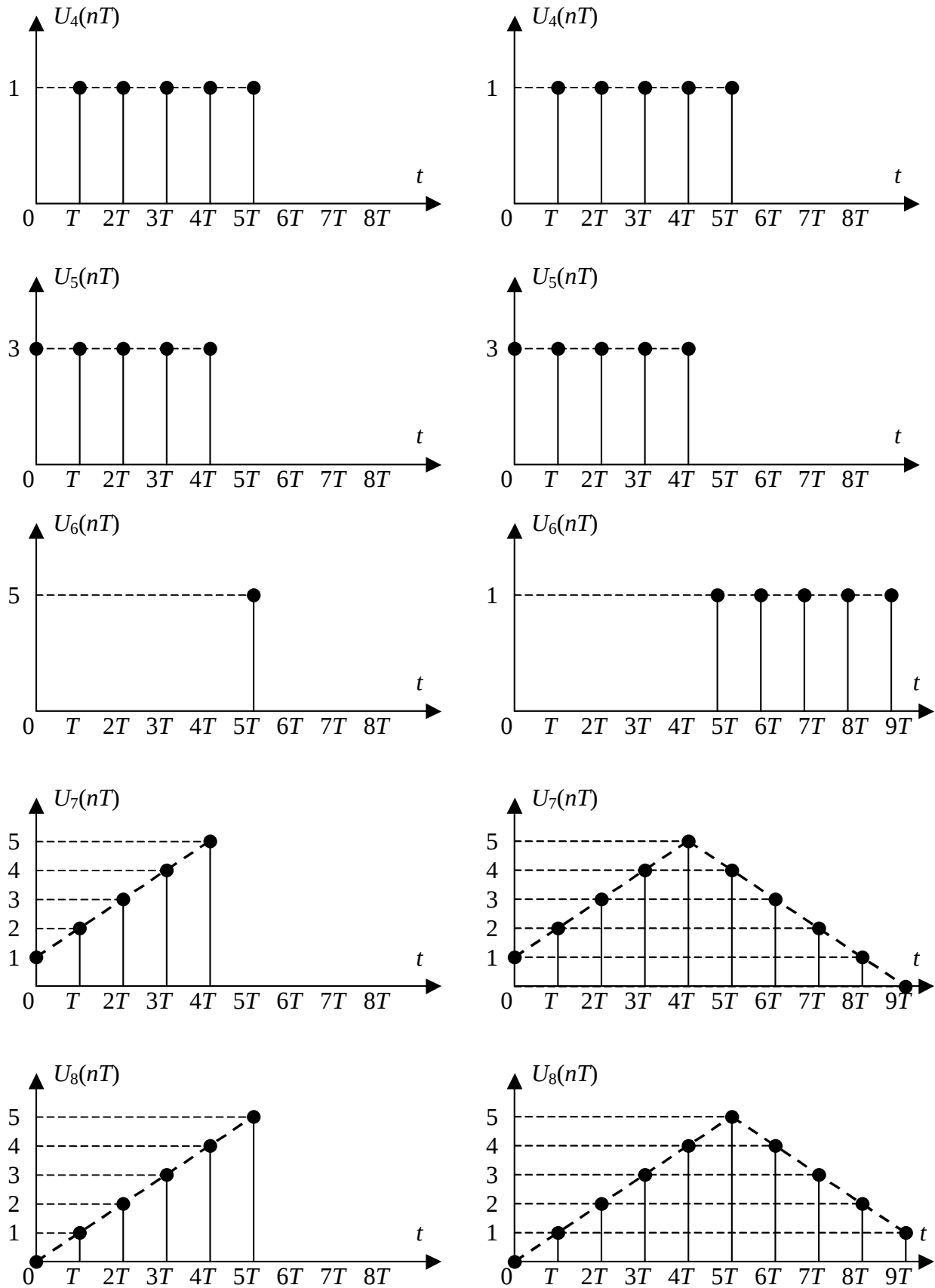


Рис. 3.39. – (продолжение)

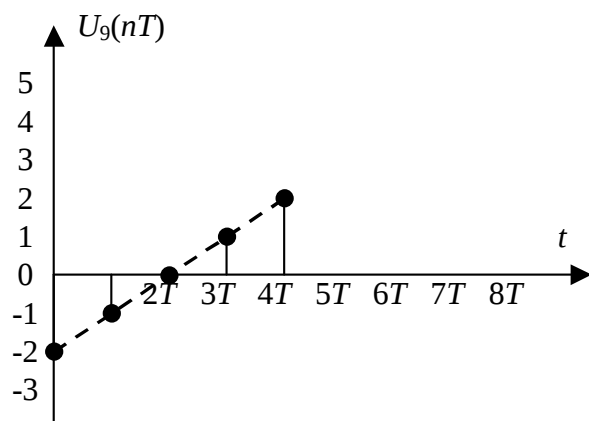


Рис. 3.39. – (окончание)

На вход устройства (рис. 3.38) в качестве сигнала $x(nT)$ подан единичный импульс (сигнал U_1) и поэтому все последующие эпюры имеют вид частных импульсных характеристик РСЦФ. В частности сигнал U_3 характеризует вид импульсной характеристики «скользящего среднего» при $M = 5$. При включении переключателя в положение 1 на выходе $y_2(nT)$ реализуется двухполярная линейная импульсная характеристика (сигнал U_9), которая характеризует фильтрующие свойства интервального дифференцирования. При включении переключателя в положение 2 на выходе $y_3(nT)$ формируются треугольная импульсная характеристика, то есть реализуется сглаживающий цифровой фильтр.

Следует отметить, что импульсные характеристики РСЦФ формируются при минимально возможном количестве арифметических операций, как правило, включающих в себя 1...2 умножения и 2...4 сложения (вычитания). В случае ортогональной обработки изображения (сепарабельность вдоль строки и кадра) элементы задержки в строчном каскаде фильтра имеют время задержки $T_э$ и $M_1T_э$, а в кадровом – $T_с$ и $M_2T_с$. При фильтрации в наклонном направлении как в строчном, так и в кадровом каскаде используются элементы задержки на время $(lT_с + mT_э)$ и $M(lT_с + mT_э)$. Здесь $T_с$ – период строки; $T_э$ – длительность элемента (и период дискретизации).

На рис. 3.40 показана структурная схема базовой ячейки РСЦФ. Рециркулятор включает в себя две схемы задержки на время $(lT_с + mT_э)$ и $M(lT_с + mT_э)$, два арифметическо-логических устройства (АЛУ1, АЛУ2) и схему управления ОЗУ1 и ОЗУ2, которые реализуют необходимую величину задержки видеосигнала.

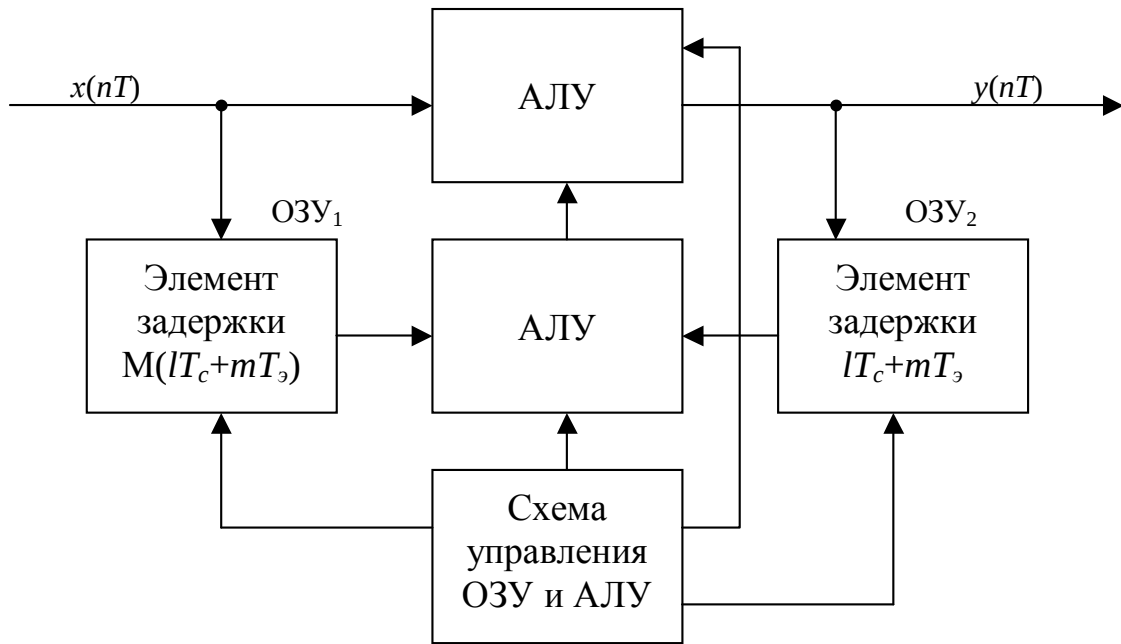


Рис. 3.40. – Структурная схема рециркулятора

При построении линии задержки на ОЗУ можно задерживать сигналы как на элементы, так и на строки. Если нет задержки на строки ($l = 0$), то лучше ставить обычные регистры. Часто при обработке изображений возникает необходимость задерживать сигнал на несколько строк. При реализации задержки на строку, необходимо задерживать сигнал как в ТВ стандарте на количество элементов в строке, причем выполнить такую задержку на регистрах уже невозможно.

При $m = 0$, $l \neq 0$ изображение фильтруется по вертикали. Если $m \neq 0$, $l \neq 0$, фильтрация будет производиться в наклонных направлениях. Величина наклона определяется конкретными значениями задержек m и l , а знак числа m определяет направление наклона фильтрующей процедуры: «+» – вправо, «-» – влево. Схему управления в режиме наклонной фильтрации можно выполнить с применением фазированных счетчиков, которые формируют адреса ОЗУ₁ и ОЗУ₂ с учетом наклона. На рис. 3.41 показана рекурсивная обработка двумерных сигналов (изображений) в неортогональных (наклонных) направлениях.

На рис. 3.41 показан общий вид двумерных импульсных характеристик для различных размеров *рекурсивной апертуры* и параметров наклона фильтрующей процедуры.

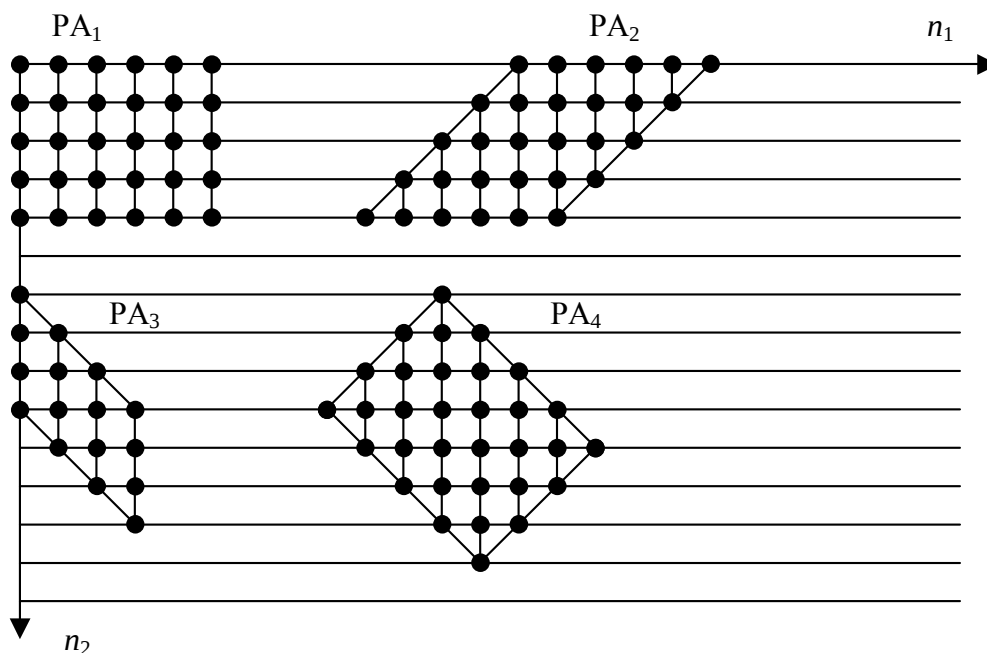


Рис. 3.41. – Рекурсивные апертуры (РА)

- $PA_1: l = 0, m = 1, M = 6 \rightarrow (T_9), (6T_9) - \text{строчный рециркулятор (СР)};$
 $l = 1, m = 0, M = 5 \rightarrow (T_9), (5T_9) - \text{кадровый рециркулятор (КР)};$
 $PA_2: l = 0, m = 0, M = 6 \rightarrow (T_9), (6T_9) - \text{СР};$
 $l = 1, m = -1, M = 5 \rightarrow (T_c - T_9), 5(T_c - T_9) - \text{КР};$
 $PA_3: l = 1, m = 1, M = 4 \rightarrow (T_c + T_9), 4(T_c + T_9) - \text{СР};$
 $l = 1, m = 0, M = 4 \rightarrow (T_c), (4T_c) - \text{КР};$
 $PA_4: l = 1, m = 1, M = 5 \rightarrow (T_c + T_9), 5(T_c + T_9) - \text{СР};$
 $l = 1, m = -1, M = 4 \rightarrow (T_c - T_9), 4(T_c - T_9) - \text{КР}.$

При $l=0, m=1$ в строчном каскаде и $m=0, l=1$ в кадровом каскаде фильтра при двумерной фильтрации реализуется прямоугольная (в пространственных направлениях) импульсная характеристика, которая ориентирована параллельно строкам раstra – первая рекурсивная апертура (PA_1). Вторая рекурсивная апертура (PA_2) реализуется последовательным включением строчного каскада фильтра с элементами задержки на T_9 и $6T_9$ и кадрового каскада с элементами задержки на $(T_c - T_9)$ и $5(T_c - T_9)$. Третья рекурсивная апертура (PA_3) реализуется аналогично при выборе времени задержки элементов в строчном каскаде $(T_c + T_9)$ и $4(T_c + T_9)$, а в кадровом каскаде T_c и $4T_c$. Четвертая рекурсивная апертура (PA_4) имеет строчные элементы задержки на время $(T_c + T_9)$ и $5(T_c + T_9)$, а кадровые элементы задержки на время $(T_c - T_9)$ и $4(T_c - T_9)$. Таким образом, выбором времени задержки узлов РСЦФ, а также коэффициентов фильтра, становится возможным реализовать широкий класс двумерных КИХ в выбранных пространственных направлениях. При использовании цифровых элементов задержки в РСЦФ, которые, как правило, реализуются с помощью ОЗУ и сдвиговых регистров, изменение времени задержки выполняется посредством коммутации выходов регистров либо изме-

нением коэффициентов деления в адресных счетчиках управляющих записью и считыванием информации ОЗУ. Произведем оценку быстродействия и аппаратурной сложности рекурсивно-сепарабельных ячеек обработки видеoinформации и оценим выигрыш по числу выполняемых арифметических операций и аппаратурными затратами по сравнению с нерекурсивными многоэлементными апертурами предназначенными для обработки видеосигнала.

Например, в случае нерекурсивной реализации пространственного фильтра, реализующего операцию «скользящее среднее», необходимо суммировать M_1M_2 отсчетов видеосигнала за один такт обработки, для этого необходимо иметь (M_1M_2-1) двухвходовых сумматоров, а также набор линий задержек по строке и кадру с (M_1-1) и (M_2-1) отводами, соответственно. Рекурсивно-сепарабельное устройство, выполняющее функцию вычисления «скользящего среднего», содержит два двухвходовых сумматора, два двухвходовых вычитателя и четыре линии задержки без отводов. Таким образом, при размере апертуры, например, 8×8 элементов выигрыш по аппаратурной сложности либо быстродействию составляет более 10 раз. При реализации операции интервального интегрирования с линейной импульсной характеристикой на базе рекурсивно-сепарабельной процедуры выигрыш по сравнению с нерекурсивным вариантом реализации достигает десятков раз, так как подобный нерекурсивный фильтр имеет в своем составе большое число умножителей, а в то время как рекурсивно-сепарабельный вариант лишь два умножителя (зачастую их можно реализовать с использованием операций сдвига чисел).

3.4. Нелинейная обработка изображений

Если в линейных процедурах обработки изображений используются только линейные операции, то в нелинейных фильтрах используются как линейные, так и нелинейные процедуры, такие как взятие модуля, возведение в квадрат, извлечение квадратного корня и т.д.

Гомоморфная фильтрация – применяется для повышения качества изображения, которое состоит из двух компонент (мультипликативная модель), т.е. $i(x, y) = i'(x, y) \cdot r(x, y)$, где $i'(x, y)$ – освещающий пучок, а $r(x, y)$ – коэффициент отражения от поверхности сцены; $i(x, y)$ – отраженный сигнал.

При гомоморфной фильтрации значительно улучшается изображение предметов, расположенных в тени.

3.4.1. Ранговая обработка изображений

Особое место среди алгоритмов нелинейной обработки изображений занимают процедуры ранговой фильтрации данных.

Ранговой обработкой данных называют преобразование набора отсчетов сигнала в *вариационный ряд* (ВР), в котором отсчеты сигнала расставляются по возрастанию значений амплитуды.

Вариационный ряд строится для каждого нового положения опорной области при развертке изображения. Опорную область также называют апертурой.

Апертура – это совокупность выделенных соседних элементов. Апертура обладает формой, размерами, «прозрачностью» и может иметь форму круга, квадрата, рамки, креста, кольца и т.д.

На рис. 3.42 приведен пример апертуры 3×3 ($x(n_1, n_2) \in A$). Таким обра-

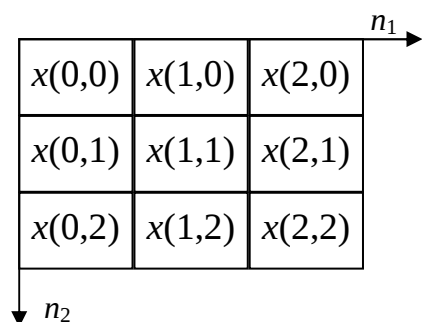


Рис. 3.42

зом, мы выделяем 9 элементов изображения при произвольном положении апертуры в изображении и по ним строим вариационный ряд, например: $\{x(n_1, n_2) \in A\} = \{3, 1, 2, 2, 1, 3, 1, 2, 4\}$, здесь вариационный ряд: $\{1, 1, 1, 2, 2, 2, 3, 3, 4\}$. Для данного примера медианное значение равно двум, а скользящее среднее равно 2,1.

Построение ВР производится по различным алгоритмам сортировки чисел (методы "камешка", "пузырька" и т. д.).

Наиболее часто используются алгоритмы экстремальной и медианной фильтрации. При экстремальной фильтрации текущий отсчет заменяется минимальным или максимальным значением сигнала из его окрестности. Экстремальные фильтры применяются для утончения или утолщения линий (например: цифр, букв, иероглифов). При медианной фильтрации текущий отсчет заменяется на значение сигнала, соответствующее центральному элементу вариационного ряда. На рис. 3.43 показан результат обработки входной последовательности линейным и нелинейным (медианным) фильтром.

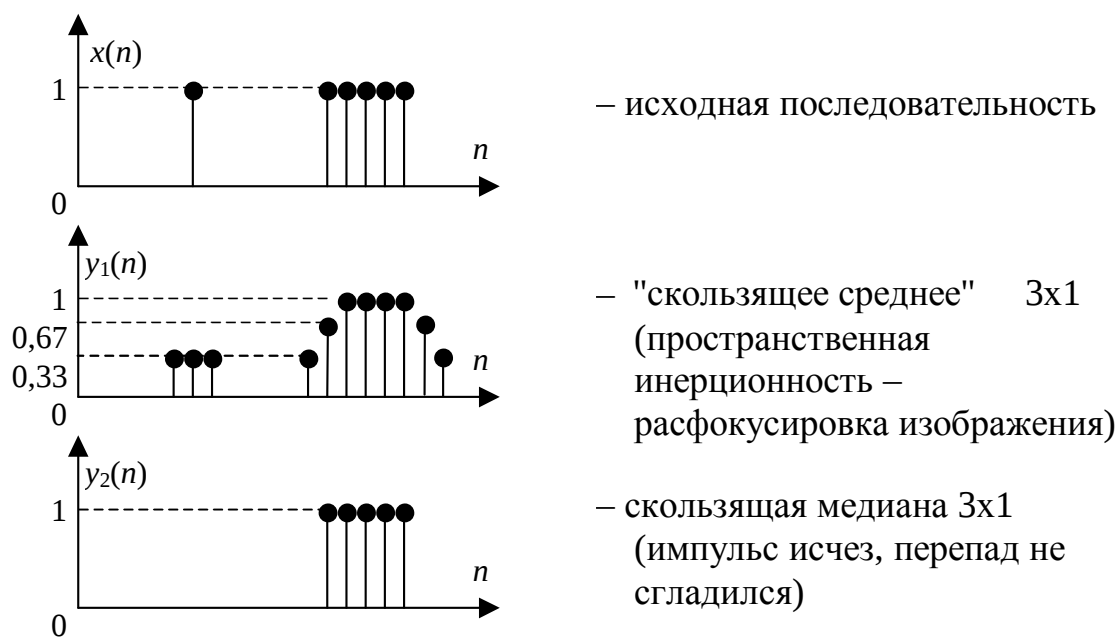


Рис. 3.43

Алгоритм двумерного медианного фильтра:

$$y(n_1, n_2) = \text{med}\{x(n_1 - i_1, n_2 - i_2)\}, \quad n_1 \geq 0, \quad n_2 \geq 0, \quad (i_1, i_2) \in A,$$

где n_1, n_2 – текущие значения координат на изображении; i_1, i_2 – координаты элементов внутри апертуры. Здесь используется апертура в виде пятиточечного креста (рис. 3.44).

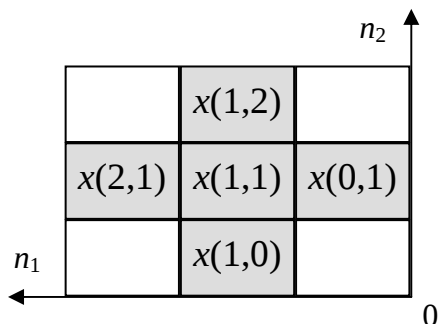


Рис. 3.44

При обработке изображений, особенно телевизионных, необходимо выделять малоразмерные детали изображения, взаимное расположение которых используется, например, при распознавании объектов и символьной информации (номеров автомобилей, вагонов, надписей на таре).

Рассмотрим работу экстремального фильтра (рис. 3.45), выполняющего операцию селекции малоразмерных объектов от протяженных фоновых образований с резкими границами.

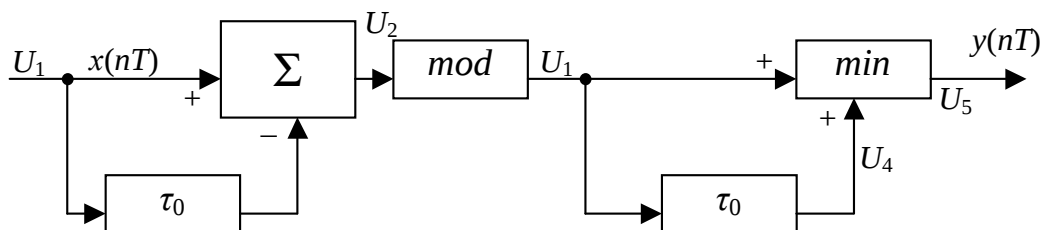


Рис. 3.45. – Структурная схема одномерного экстремального фильтра

Здесь $\tau_0 = NT$ – задержка, равная нескольким тактам (при цифровом построении).

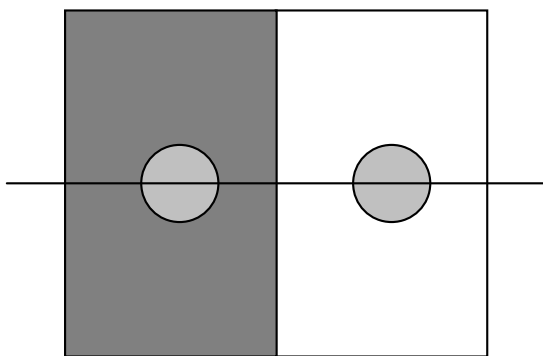


Рис. 3.46

Светлая деталь на темном фоне (самолет, ракета на фоне Земли) и темная деталь на светлом фоне (тот же самолет на фоне облаков) (рис. 3.46). Здесь объект меняет свой контраст с положительного на отрицательный. Суть нелинейной обработки заключается во взятии первой производной в строчном направлении. При этом происходит «расщепление» сигнала от объекта – (U_3) (рис. 3.47).

После обработки в блоке вычисления $\min$ происходит выделение сигнала от объекта как положительного, так и отрицательного контраста, причем объект отрицательного контраста дает положительный сигнал.

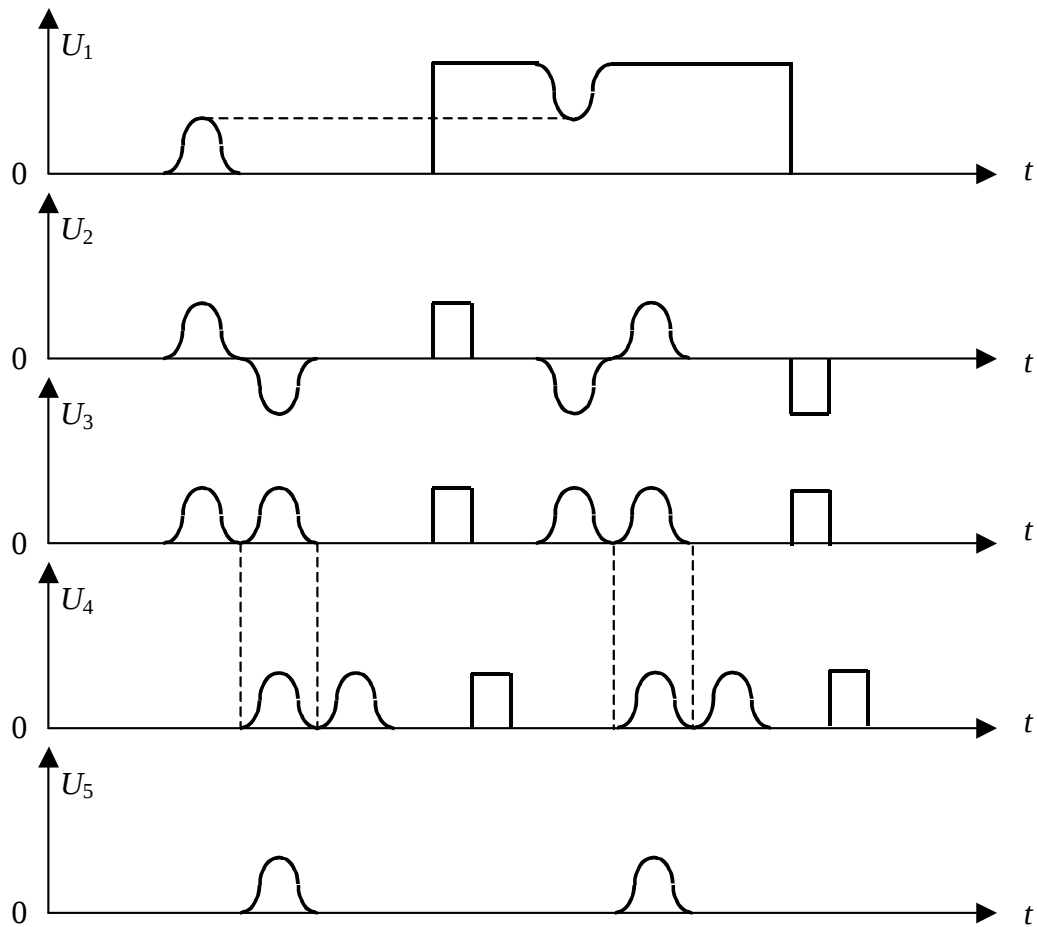


Рис. 3.47

Здесь происходит хорошее выделение сигнала от малоразмерной детали на фоне с резкими перепадами яркости (сигнал от перепада в два раза больше, чем от малоразмерного объекта, но тем не менее он полностью подавлен). Линейная обработка таких результатов не дает. При использовании, например, лапласианов или двумерного дифференцирования получается большой паразитный сигнал от перепада яркости в изображении.

Недостаток одномерного фильтра: в этом изображении «полосы» выделяются так же успешно, как и малоразмерный объект, т.к. в этом случае не используется информация о том, что «полосы» имеют протяженность в вертикальном направлении (рис. 3.48). Рассмотренный экстремальный фильтр принципиально одномерный. В нем нет обработки в кадровом направлении.

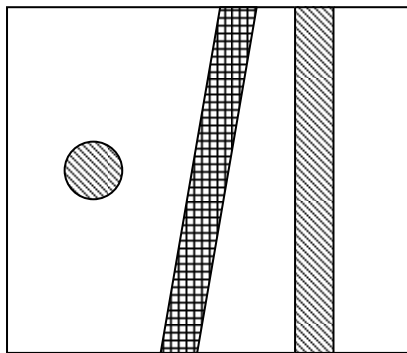


Рис. 3.48

Апертура фильтра $(2N+1)$ приведена на рис. 3.49. Здесь $\tau_0 = NT$ — равно длительности (размеру) объекта в строчном направлении.

Для устранения указанного выше недостатка

фильтра (по причине его одномерности) используются двумерные маски (апертуры), например, незаполненная маска 7x7 (рис. 3.50).

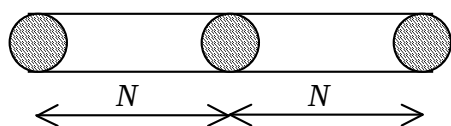


Рис. 3.49

Незаполненные маски используются для уменьшения операций в фильтре. Размеры маски должны быть по крайней мере в два раза больше размера выделяемого малоразмерного объекта. Данная маска почти изотропная (результаты обработки одинаковы во всех направлениях) и с ее помощью мы лишь определяем элементы, которые будут участвовать в обработке, а не саму обработку. В процессе обработки происходит вычисление локальных разностей между центральным и периферийными элементами (для рис.3.50 – 8 первых разностей). На рис. 3.51 приведена структурная схема фильтра, использующего незаполненную маску с рис. 3.50.

			(3, 6)			
	(5, 5)				(1, 5)	
(6, 3)			(3, 3)			(0, 3)
	(5, 1)				(1, 1)	
			(3, 0)			

Рис. 3.50

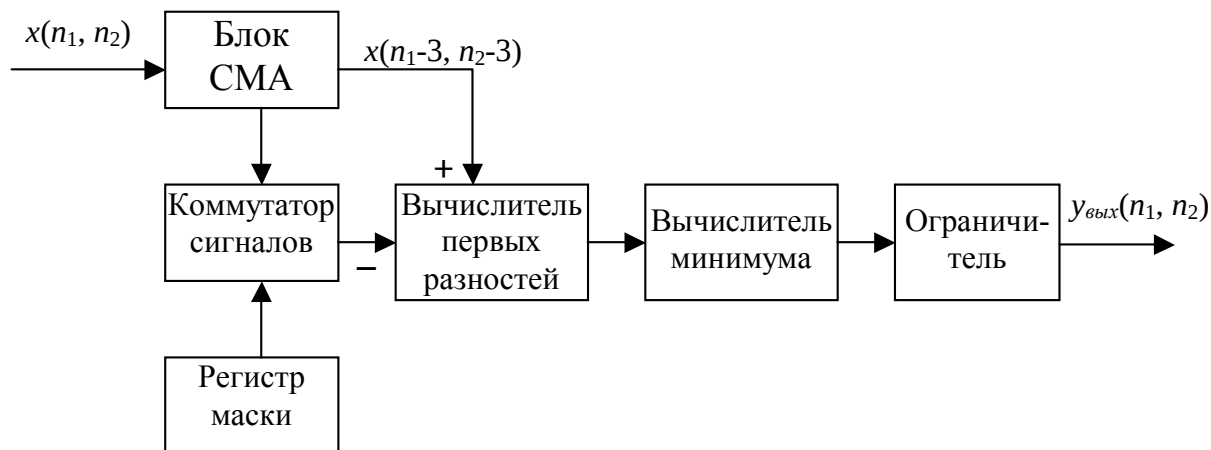


Рис. 3.51. – Двумерный экстремальный фильтр с апертурой 7x7 элементов (нелинейный нерекурсивный фильтр)

Нелинейное разностное уравнение имеет вид

$$y(n_1, n_2) = \min \{ x(n_1 - 3, n_2 - 3) - x(n_1 - 3, n_2); x(n_1 - 3, n_2 - 3) - x(n_1 - 5, n_2 - 1); \\ x(n_1 - 3, n_2 - 3) - x(n_1 - 6, n_2 - 3); \dots; x(n_1 - 3, n_2 - 3) - x(n_1 - 1, n_2 - 1) \}.$$

Всего здесь используется 8 локальных первых разностей и после вычисления минимальной из них выполняется операция ограничения.

$$y_{\text{блх}}(n_1, n_2) = f[y(n_1, n_2)],$$

$$\text{где } f(x) = \begin{cases} x, & \text{если } x \geq 0, \\ 0, & \text{иначе.} \end{cases}$$

Регистр маски выдает адреса лишь тех элементов, которые участвуют в обработке. Любой малоразмерный объект есть множество точек в некоторой компактной области (это может быть круг, прямоугольник, крест, ромб и т.д.). Вокруг центрального элемента есть «тормозная зона». Если размер этой зоны в два раза больше размеров объекта, то объект выделяется без искажений. Для фона, т. к. он протяженный и в пределах апертуры изменяется слабо, алгоритм обладает разрушительным воздействием. Для фона с резким перепадом всегда найдется первая разность, которая меньше или равна нулю, и она пройдет на выход. Таким образом, сигнал от фона после обработки всегда будет ниже уровня ограничения и не пройдет на выход.

Схему на приведенном выше рисунке можно существенно упростить (рис. 3.52).

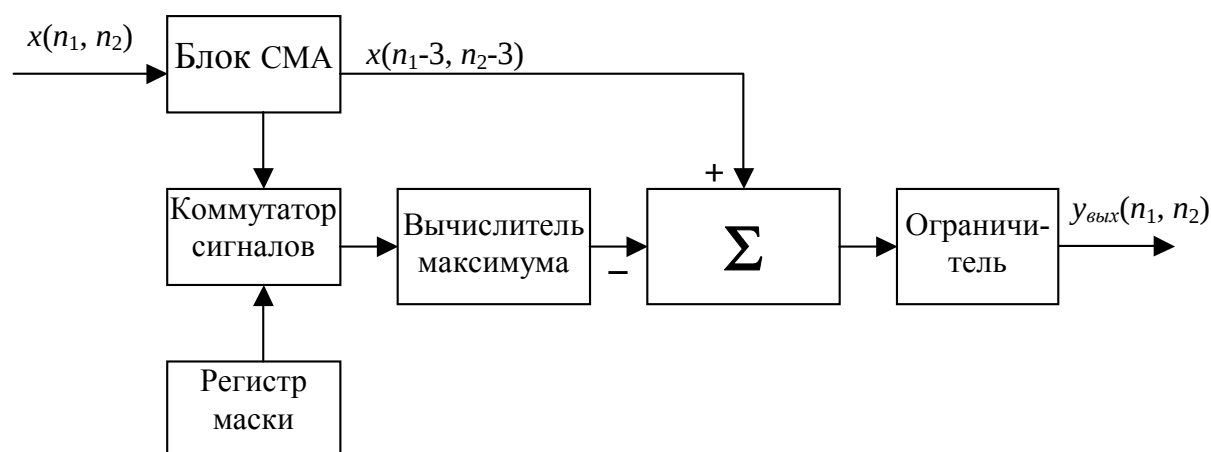


Рис. 3.52. – Двумерный экстремальный фильтр

Нелинейное разностное уравнение такого фильтра имеет вид

$$y(n_1, n_2) = x(n_1 - 3, n_2 - 3) - \max\{x(n_1 - 3, n_2); x(n_1 - 5, n_2 - 5); x(n_1 - 6, n_2 - 3); \\ \dots; x(n_1 - 1, n_2 - 1)\},$$

$$y_{\text{блх}}(n_1, n_2) = f[y(n_1, n_2)],$$

где $f[\cdot]$ – функция ограничения снизу, т.е. все отрицательные выходные значения становятся нулевыми.

Вынесем центральный элемент $x(n_1-3, n_2-3)$ из операции $\max\{\cdot\}$. Также вынесем знак минуса от остальных элементов, тогда операция $\max\{\cdot\}$ станет $\min\{\cdot\}$. Т.е. выражения $\min\{-2; -3\} = -3$ и $-\max\{2; 3\} = -3$ эквивалентны.

Выше был описан так называемый *метод минимума первых разностей*. При всей своей простоте он имеет серьезный недостаток – выделяет только

объекты положительного контраста, а объекты отрицательного контраста подавляются из-за наличия функции ограничения снизу.

Для выделения объектов с контрастом любой полярности можно использовать *метод минимума модулей первых разностей*. Как можно догадаться, структурную схему данного алгоритма можно получить, если на рис. 3.51 перед оператором минимума включить оператор взятия модуля вычисляемых первых разностей. Здесь уже нет необходимости в функции ограничения снизу, поскольку мы оперируем только с положительными величинами, поэтому ограничитель со структурной схемы можно убрать. Разностное уравнение этого алгоритма обработки имеет вид

$$y_{\text{вых}}(n_1, n_2) = \min\{|x(n_1 - 3, n_2 - 3) - x(n_1 - 3, n_2)|; \dots; |x(n_1 - 3, n_2 - 3) - x(n_1 - 1, n_2 - 1)|\}.$$

Однако и этот алгоритм не лишен недостатков – корректно выделяются только объекты с резкими перепадами сигнала (см. рис. 3.46). Если объект имеет «плавные» края, например, двумерная гауссоида, и его размеры менее чем в два раза меньше размеров маски, то происходит дробление выделяемого объекта по его краям (если бы круглый объект на рис. 3.46 имел плавные края, то на выходе вместо круга была бы получена восьмиконечная звезда) из-за наличия операции взятия модуля.

Указанных выше недостатков лишен третий алгоритм – *модифицированный метод минимума первых разностей*. Его разностное уравнение имеет вид

$$y_{\text{вых}}(n_1, n_2) = f[x_c - \min\{x_i\}] + f[\max\{x_i\} - x_c],$$

где x_c – центральный элемент маски, т.е. $x(3, 3)$ (см. рис. 3.50), а $\{x_i\}$ – остальные элементы маски (они выделены серым цветом на рис. 3.50).

3.5. Двумерная децимация данных

Для решения задач обнаружения объектов в телевизионном изображении и измерения параметров объектов (координат, размеров, яркости и т.д.) необходимо обрабатывать большой поток видеоданных в реальном масштабе времени. Большая часть видеоинформации является малоинформативной, поэтому до реализации алгоритмов обнаружения объектов целесообразно применять операцию сжатия (кодирования) данных. Под сжатием видеоданных понимают либо уменьшение полосы частот видеосигнала, либо уменьшение времени передачи сигнала при неизменной полосе частот (т.е. в той же полосе либо за то же время передается больше данных). Критерием оценки качества сжатия видеоданных является обеспечение заданных характеристик обнаружения объектов и точности измерения их параметров.

К достаточно простым и эффективным методам сжатия данных относятся прореживание отсчетов сигнала и уменьшение частоты дискретизации по-

сле предварительной низкочастотной фильтрации, исключаяющей эффекты наложения спектров.

Системы уменьшения частоты дискретизации (*децимации*) относятся к классу нисходящих дискретных систем. Для двумерных сигналов необходимо ввести понятие коэффициента децимации в горизонтальном (m_1) и вертикальном (m_2) направлениях. В общем случае m_1 и m_2 – произвольные числа, но тогда выполнение операции двумерной децимации дискретного изображения $x(n_1, n_2)$ существенно затрудняется, так как необходимо для вычисления отсчетов обеспечивать двумерную интерполяцию сигнала. Рассмотрим случай двумерной децимации сигнала, когда m_1 и m_2 – положительные целые числа. При $m_1=m_2=1$ выходной сигнал устройства децимации полностью совпадает с входным сигналом. Выходной сигнал устройства прореживания отсчетов сигнала без фильтрации записывается в виде операции пространственного маскирования массива

$$y_n(m_1 n'_1, m_2 n'_2) = \delta(m_1 n'_1, m_2 n'_2) x(n_1, n_2).$$

здесь функция $\delta(m_1 n'_1, m_2 n'_2)$ осуществляет вторичную дискретизацию массива данных с заданными периодами в горизонтальном и вертикальном направлениях (n'_1 и n'_2 – целые положительные числа с шагом изменения равным единице). Операция уменьшения частот пространственной дискретизации двумерного сигнала с предварительной фильтрацией данных записывается в виде

$$y_d(m_1 n'_1, m_2 n'_2) = \delta(m_1 n'_1, m_2 n'_2) [h(n_1, n_2) ** x(n_1, n_2)],$$

где $h(n_1, n_2)$ – двумерная импульсная характеристика фильтра, а $**$ – обозначает операцию двумерной свертки сигнала и импульсной характеристики фильтра. При использовании в качестве фильтров в устройстве децимации сигнала каскадно включенных разделимых (сепарабельных) ячеек, возможно выполнение операций прореживания отсчетов сигнала поэтапно, например, вначале в горизонтальном, а затем в вертикальном, а затем в вертикальном направлениях. Операция *раздельной двумерной децимации с фильтрацией данных* записывается в виде

$$y(m_1 n'_1, m_2 n'_2) = \delta(m_2 n'_2) \{ h(n_2) ** \delta(m_1 n'_1) \times [h(n_1) ** x(n_1, n_2)] \},$$

здесь $\delta(m_1 n'_1)$ и $\delta(m_2 n'_2)$ – дискретизирующие функции, в горизонтальном и вертикальном направлениях, соответственно, а $h(n_1)$ и $h(n_2)$ – импульсные характеристики каскадов разделимого двумерного фильтра.

При использовании нерекурсивных алгоритмов фильтрации в устройстве децимации вычисление выходных отсчетов сигнала производится только в узлах вторичной дискретизации ($m_1 n'_1, m_2 n'_2$), а при использовании рекуррентных алгоритмов вычисление приходится выполнять в узлах первичной дискретизации (n_1, n_2), что снижает эффективность их использования в подобных устройствах.

Сжатие видеоданных осуществляется с целью уменьшения массива данных, которые поступают для обработки на вычислитель. Кроме того, при двумерной децимации видеоданных возможна реализация алгоритмов, которые наилучшим образом подготавливают информацию для решения задачи обнаружения объектов либо измерения их параметров. Рассмотрим характерные алгоритмы обработки видеоинформации с предварительной децимацией данных.

Для работы в реальном масштабе времени обработка видеоинформации должна выполняться на время, как правило, не превышающее длительности телевизионного поля (для вещательного стандарта разложения – 0,02 с). Объем информации, который может быть обработан за указанное время, ограничен. При производительности вычислителя, равной к примеру 10^6 операций в секунду, реализация алгоритмов обнаружения деталей изображения либо измерения их координат требует времени на обработку каждого из отсчетов порядка $20 \cdot 10^{-6}$ с. Таким образом, размер обрабатываемого массива видеоданных в реальном масштабе времени не может превышать 32×32 элементов разложения. Размеры деталей в телевизионном изображении, как правило, превышают указанный размер фрагмента изображения, а при обнаружении мало-размерных деталей изображения необходимо анализировать все поле зрения телевизионного датчика. В этом случае, для сохранения заданной величины быстродействия, необходимо произвести операцию децимации с предварительной двумерной фильтрацией данных.

Для аппаратной реализации устройства двумерной децимации видеоинформации можно успешно использовать рекурсивно-сепарабельные ячейки, описанные в п. 3.3, которые включаются каскадно. Операция прореживания сглаженной информации выполняется по строке с шагом m_1 элементов, а по кадру с шагом m_2 строк. Таким образом число элементов в обрабатываемом изображении становится $m_1 \times m_2$ раз меньше, чем в исходном и требования к производительности вычислителя существенно снижаются.

Произведем оценку качества выполнения задач обнаружения объектов в изображении и измерения их координат в телевизионной измерительной системе (ТИС) со сжатием видеоданных при использовании наиболее распространенных алгоритмов обработки видеоинформации.

При решении задачи выделения подвижных областей в телевизионном изображении часто используется межкадровый алгоритм вида

$$Q(n_1, n_2, n_3) = S(n_1, n_2, n_3) - S(n_1, n_2, n_3 - 1) + (1 - 1/A) Q(n_1, n_2, n_3 - 1),$$

здесь $S(n_1, n_2, n_3)$ и $Q(n_1, n_2, n_3)$ – последовательности во времени входных и выходных изображений, соответственно; A – параметр фильтрации; n_1 и n_2 – номера элементов и строк в изображении, соответственно; n_3 – операнд дискретного времени (номер текущего кадра). В случае введения операции децимации данных до межкадровой обработки на его вход подается сжатое в $m_1 m_2$ раз входное изображение и, соответственно, выходное изображение содержит

также меньшее число отсчетов. Приведенное выше выражение для децимированного потока данных можно записать в том же виде, приняв $n_1 = n'_1 m_1$ и $n_2 = n'_2 m_2$, где m_1 и m_2 – коэффициенты децимации по строке и кадру, соответственно.

Для выделения областей в изображении с изменяющейся амплитудой перед межкадровой обработкой достаточно выполнить операцию «скользящего среднего» над входными данными и произвести их прореживание. Выделение подвижных областей в изображении при таком кодировании отсчетов сигнала возможно только при их переходе из одной децимирующей ячейки в другую, что не позволяет выделять малоподвижные области в телевизионном изображении за заданный интервал времени. Наиболее эффективным представляется алгоритм сжатия информации с нормированием в соответствии с формулой «центра масс», так как в этом случае производится наиболее полное кодирование координатной информации о фрагментах изображения.

На рис. 3.53 приведена структурная схема устройства межкадровой обработки с предварительным кодированием (сжатием) видеоданных. На базе

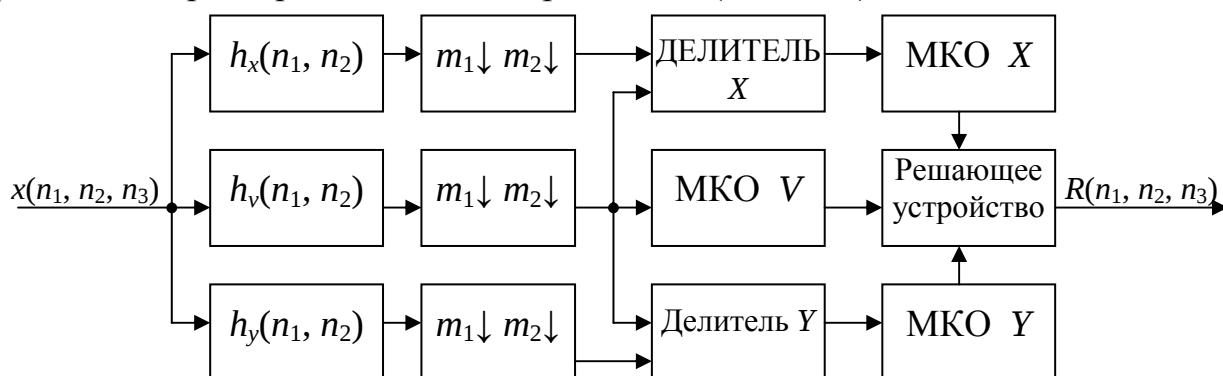


Рис. 3.53. – Структурная схема устройства межкадровой обработки (МКО) с предварительным сжатием данных

устройств такого типа можно строить эффективные обнаружители подвижных объектов на подвижном фоне либо селектировать объекты с заданной скоростью. Поясним принцип работы данного устройства. Цифровой видеосигнал $x(n_1, n_2, n_3)$ поступает на входы трех фильтров.

Размеры апертуры фильтров выбираются исходя из размеров объема обнаружения и размеров изображений объектов. С выходов фильтров, которые могут быть построены на основе рекурсивно-сепарабельных ячеек, сигналы прореживаются по строке и кадру блоками $(m_1 \downarrow m_2 \downarrow)$, нормируются делителями X и Y , соответственно, и подаются на входы блоков межкадровой обработки (МКО) данных, работающих в соответствии с приведенным выше алгоритмом (см. предыдущую формулу). Таким образом, на входы блоков МКО X МКО Y подаются значения сигналов рассогласований по строке $\Delta X(m_1 n'_1, m_2 n'_2, n_3)$ и кадру $\Delta Y(m_1 n'_1, m_2 n'_2, n_3)$, соответственно, вычисленные по формуле «центра масс», для каждой из децимируемых ячеек. Кроме того, сумма отсчетов сигналов в каждой из ячеек, которая является нормирующей величиной для вы-

числения сигналов рассогласования $V(m_1n'_1, m_2n'_2, n_3)$, также подается на блок МКО V . Далее сигналы с выходов блоков МКО поступают на решающее устройство, которое по входной информации может эффективно формировать сигнал обнаружения подвижных объектов в телевизионном изображении на фоне подвижных помех $R(n_1, n_2, n_3)$. Например, предположим, что в поле зрения телевизионного датчика находится L неподвижных помеховых образований, K – помеховых образований, движущихся с одной скоростью, и один подвижный объект, который следует обнаружить. С выхода децимирующих фильтров (при некоторой идеализации процесса обнаружения) поступают сигналы $\Delta X(m_1n'_1, m_2n'_2, n_3)$, $\Delta Y(m_1n'_1, m_2n'_2, n_3)$ и $V(m_1n'_1, m_2n'_2, n_3)$, являющиеся постоянными во времени для, как минимум, L ячеек, в которых находятся сигналы от неподвижных объектов.

После межкадровой обработки информации от этих ячеек выходной сигнал с блоков МКО становится равным нулю, то есть неподвижные помеховые образования будут подавлены и они в дальнейшем при обнаружении не участвуют. Подобная информация от, как минимум, K ячеек после межкадровой обработки сигналов $\Delta X(m_1n'_1, m_2n'_2, n_3)$ и $\Delta Y(m_1n'_1, m_2n'_2, n_3)$ должна представлять равные (близкие) между собой величины, так как эта группа помеховых образований по условию движется с одинаковой скоростью. Таким образом, при построении функции распределения выходных сигналов блоков МКО можно обеспечить попадание сигналов от группы из K помеховых образований в один интервал ее разбиения. Анализируя распределение сигналов с выходов блоков МКО, обрабатывающих массивы сигналов рассогласования по строке и кадру, можно с достаточной степенью достоверности обнаружить объекты с заданной скоростью либо объекты, движущиеся со скоростью отличной от скорости помеховых образований.

Для оценки характеристик качества обнаружения подвижных объектов в телевизионном изображении на фоне белого шума можно произвести вычислительный эксперимент по моделированию работы устройства, приведенного на рис. 3.53. В этом случае на вход модели устройства подается сигнал от изображения, содержащего подвижный объект гауссоидальной формы при разных скоростях его движения: $v = 0,1; 0,5; 1$ элемент разложения за кадр (эл/к). Размер объекта принят равным $r_{0,606} = 1$; два элемента разложения; отношение сигнал/шум на входе устройства $\psi_{\text{вх}} = 10$.

На рис. 3.54 приведены графики, иллюстрирующие отношение сигнал/шум $\psi_{\text{вых}}$ на выходе канала межкадровой обработки сигнала $V(m_1n'_1, m_2n'_2, n_3)$ при постоянной фильтрации ($A = 8$), коэффициентах децимации по строке и кадру $m_1 = m_2 = m_d$ и размере апертуры фильтра равном коэффициенту децимации $M = m_d$. При обработке сигнала в этом случае происходит увеличение $\psi_{\text{вых}}$ при значениях коэффициента $m_d \leq 2r_{0,606}$. По результатам моделирования можно сделать вывод о целесообразности сжатия видеoinформации до межкадровой обработки при решении задачи обнаружения подвижных

объектов в телевизионном изображении. Коэффициент, сжатия видеоданных m_s , в данном случае равен $m_s = m_d^2$ и может достигать значений от 5 до 20 (по количеству обрабатываемых отсчетов), и при этом не происходит ухудшения характеристик обнаружения.

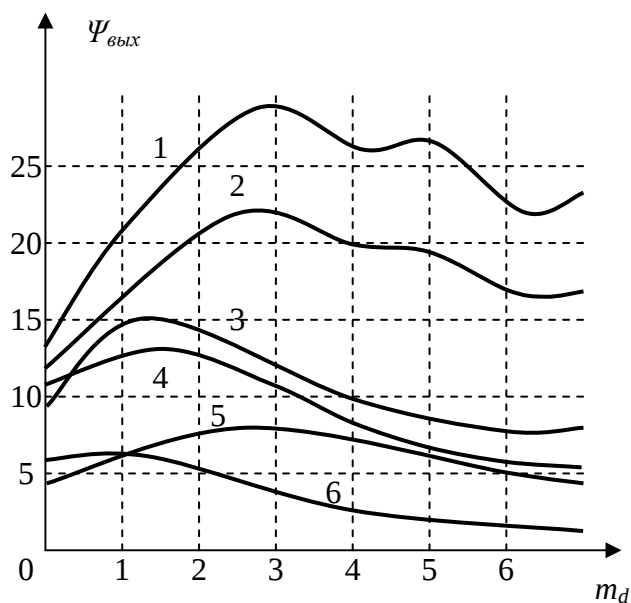


Рис. 3.54. — Зависимости отношений сигнал/шум на выходе канала межкадровой обработки сигнала от коэффициента децимации данных:

1 — $v = 1$ эл/к;	$r_{0,606} = 2$ эл;
2 — $v = 0,5$ эл/к;	$r_{0,606} = 2$ эл;
3 — $v = 1$ эл/к;	$r_{0,606} = 1$ эл;
4 — $v = 0,5$ эл/к;	$r_{0,606} = 1$ эл;
5 — $v = 0,1$ эл/к;	$r_{0,606} = 2$ эл;
6 — $v = 0,1$ эл/к;	$r_{0,606} = 1$ эл.

На рис. 3.55 приведена структурная схема измерителя координат объекта с предварительным сжатием видеоданных. В качестве децимирующего фильтра, в данном случае, используется устройство вычисляющее «скользящее среднее» в апертуре размером $M \times M$ элементов $h_v(n_1, n_2)$.

После двумерной децимации выходного сигнала фильтра, с равными коэффициентами по строке и кадру ($m_1 = m_2 = m_d$), сжатый массив сигнала подается на вход дискриминатора, вычисляющего координаты объекта по строке $\Delta X(N_1, N_2)$ и по кадру $\Delta Y(N_1, N_2)$ в заданной точке раstra (N_1, N_2). Координаты объекта вычисляются в соответствии с выражениями, описывающими алгоритм измерения «центра тяжести» двумерного распределения сигнала. Таким образом, появляется возможность при выполнении определенного количества арифметических

операций в единицу времени оценивать координаты детали изображения на порядок большей площади, чем без операции децимации видеосигнала.

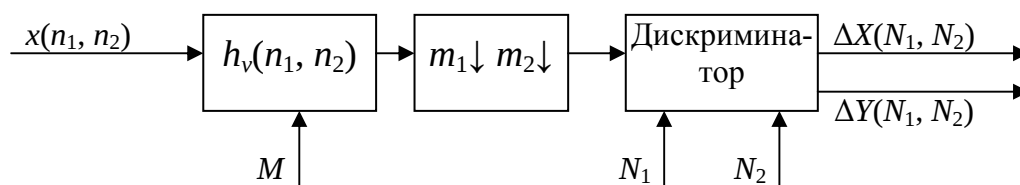


Рис. 3.55. — Структурная схема измерителя координат объекта с предварительным сжатием данных

Приведенные результаты по оценке эффективности операции децимации видеоданных в ТИС показывают, что при больших отношениях сигнал/шум $\Psi_{вых} = \infty$ существует минимум ошибки измерения координат гауссоидальной

детали изображения по децимированной информации при соответствующем размере M сглаживающей апертуры.

На рис. 3.56 приведены графики среднеквадратичной ошибки σ_x , характеризующие точность измерения координат объекта по строке, при различных параметрах устройства децимации сигнала. Минимальная ошибка измерения координат объекта реализуется при размере апертуры $M = m_d + r_{0,606}$. Для отношения сигнал/шум на входе устройства $\psi_{\text{вх}} = 20$ ошибка измерения координат при различных коэффициентах децимации m_d наименьшая по сво-

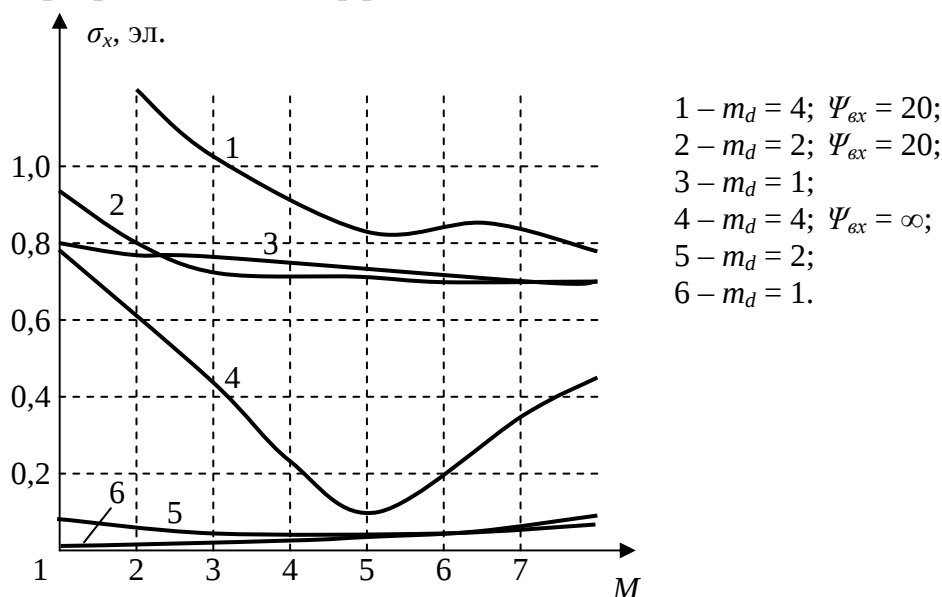


Рис. 3.56. – Зависимости средней квадратической ошибки измерения координат объекта от параметров устройства сжатия видеоданных ($r_{0,606} = 1$ эл.)

им значениям также при размере апертуры $M = m_d + r_{0,606}$. Таким образом, при правильном выборе параметров устройства децимации, может быть реализована высокая точность измерения координат объектов при уменьшении требуемой производительности вычислителя на порядок. Использование предлагаемых и более сложных алгоритмов двумерной децимации видеoinформации в ТИС позволяет улучшить их технические характеристики, а также снизить их сложность и стоимость.

4. СПЕЦИАЛИЗИРОВАННЫЕ УСТРОЙСТВА ДЛЯ ЦИФРОВОЙ ФИЛЬТРАЦИИ СИГНАЛОВ

В предыдущих разделах основное внимание уделялось вопросам анализа и синтеза алгоритмов цифровой обработки сигналов. В данном разделе будут рассмотрены специализированные цифровые устройства, предназначенные для ускоренного выполнения алгоритмов обработки и для построения систем обработки в реальном масштабе времени [5].

Важную роль в достижении максимума отношения качества системы обработки к ее стоимости играют методы распараллеливания, временного разделения и поточной обработки. Прежде всего, в данном разделе будут рассмотрены различные способы построения цифровых фильтров, в том числе прямая и каскадная формы КИХ-фильтров, прямая, каскадная и параллельная формы БИХ-фильтров. После этого будет приведено несколько практических примеров построения цифровых специализированных систем, включая реализацию цифровых фильтров.

4.1. Аппаратурное построение КИХ-фильтра прямой формы

Рассмотрим способы построения КИХ-фильтра прямой формы, схематично изображенного на рис. 4.1.

Найдем, прежде всего, основную схему управления, пригодную для построения фильтров с различным уровнем параллелизма, причем параллелизм будет заключаться не только в увеличении количества арифметических устройств, но и в параллельной работе блоков памяти.

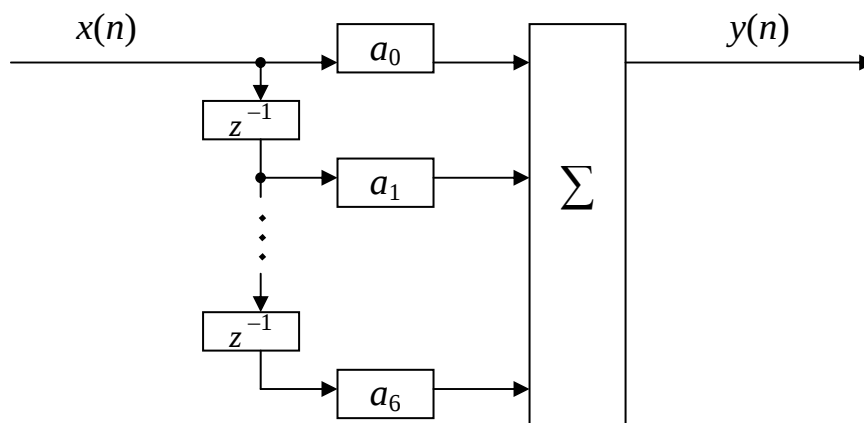
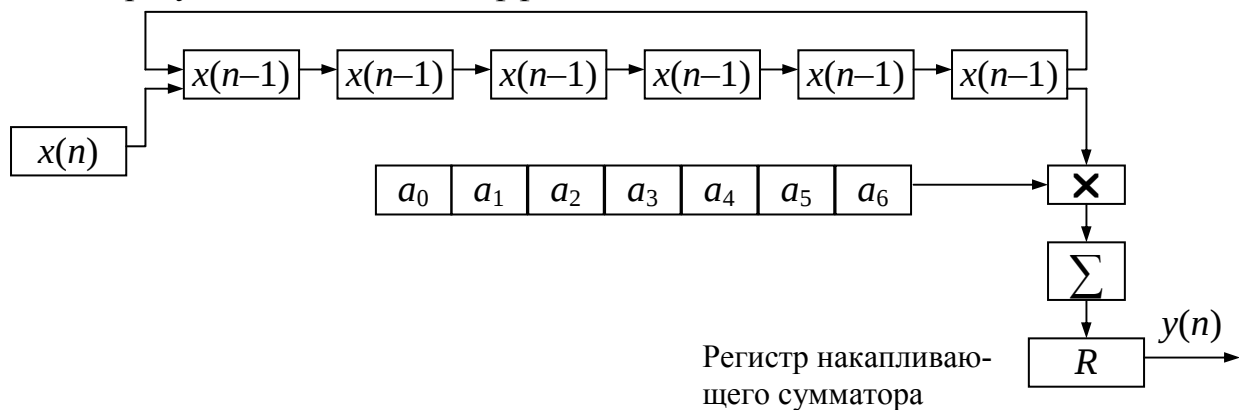


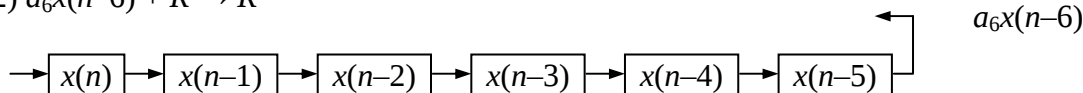
Рис. 4.1. – Прямая форма КИХ-фильтра 7-го порядка

На рис. 4.2 изображена простая блок-схема построения КИХ-фильтра прямой формы (рис. 4.1) из единственного арифметического устройства (содержащего умножитель и сумматор), регистра сдвига для хранения промежуточных результатов и ПЗУ коэффициентов.

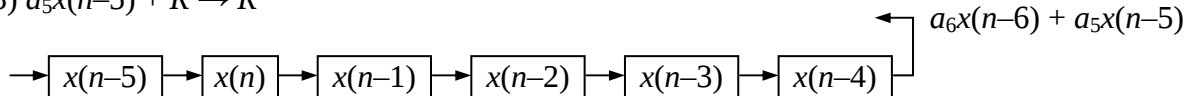


1) Очистить R

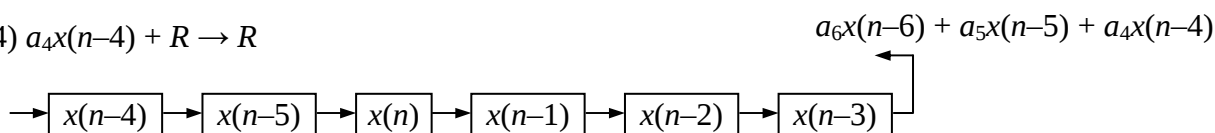
2) $a_6x(n-6) + R \rightarrow R$



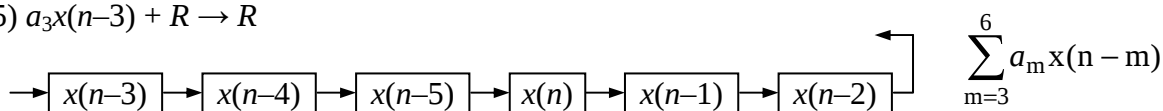
3) $a_5x(n-5) + R \rightarrow R$



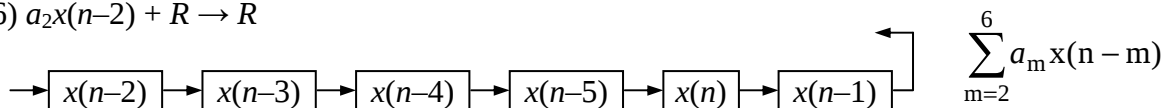
4) $a_4x(n-4) + R \rightarrow R$



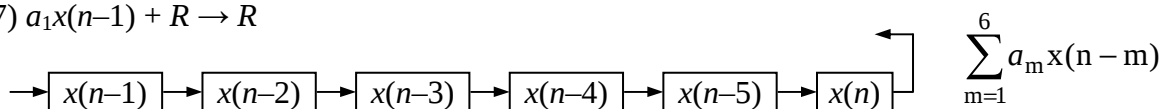
5) $a_3x(n-3) + R \rightarrow R$



6) $a_2x(n-2) + R \rightarrow R$



7) $a_1x(n-1) + R \rightarrow R$



8) $a_0x(n) + R \rightarrow R$

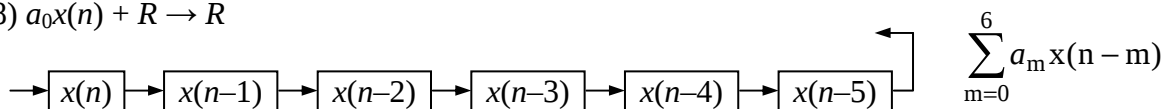


Рис. 4.2. – Структурная схема, последовательность обработки и промежуточные результаты для прямой формы КИХ-фильтра

Важным моментом является характер управления регистром сдвига. Каждый выходной отсчет вычисляется путем накопления последовательных произведений, образующихся на выходе умножителя в процессе кругового сдвига регистра. Новый входной отсчет $x(n)$ вводится в регистр одновременно с выталкиванием с противоположного конца регистра отсчета $x(n-6)$, сопровождающимся вычислением произведения $a_6x(n-6)$. При каждой итерации, как это видно из схемы вычислений на рис. 4.2, производится круговой сдвиг регистра на один отсчет. После вычисления значения $y(n)$ оно выводится; одновременно с этим очищается накопитель, после чего начинается следующий главный цикл.

Заметим, что структура памяти на сдвиговом регистре, используемой в схеме на рис. 4.2, еще не была определена. Прежде всего, необходимо выяснить, на каких микросхемах ее можно собирать. Оказывается, что для этой цели хорошо подходят большие интегральные схемы с МДП-регистрами сдвига. В настоящее время на одном кристалле могут размещаться несколько тысяч последовательно соединенных разрядов памяти. Поэтому для многих КИХ-фильтров вся память нужного объема размещается в одном корпусе, но при условии, что данные вводятся и подаются на арифметическое устройство последовательно. Конечно, было бы интересно построить КИХ-фильтр прямой формы с использованием именно такой памяти, состоящей из малого числа элементов. Простейшая схема подобного типа приведена на рис. 4.3.

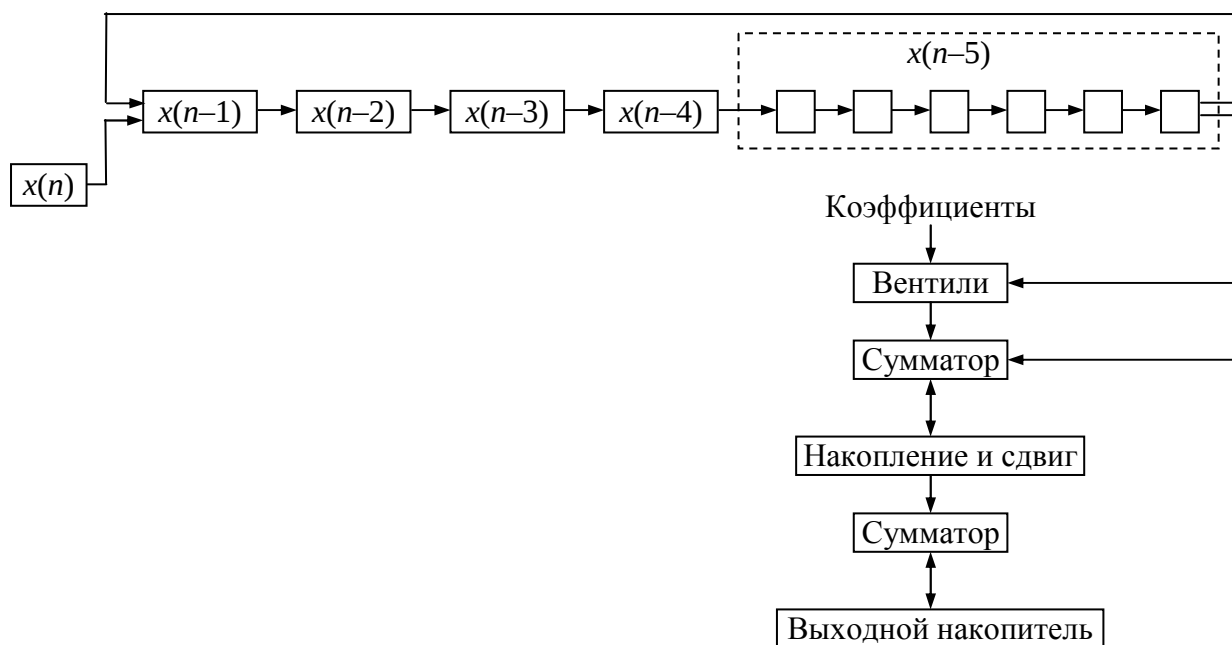


Рис. 4.3. – Построение КИХ-фильтра 6-го порядка с использованием алгоритма Бута

В этой схеме каждое произведение получается с помощью последовательности логических умножений коэффициента на разряды перемножаемого отсчета (в данном случае $x(n-6)$) и суммирования этих произведений. Все от-

счеты в схеме на рис. 4.3 хранятся в последовательном виде, хотя это показано только для отсчета $x(n-6)$.

Следует напомнить, что при выполнении умножения, основанного на сдвигах и сложениях, вид алгоритма зависит от знаков сомножителей. Так, при использовании прямого кода сложение будет обычным, а знак произведения будет определяться знаками отсчета и коэффициента. При умножении с использованием сдвигов и сложений этот подход довольно эффективен; однако заключительное накопление всех произведений при этом усложняется, так как для сложения чисел в прямом коде необходимы сумматор-вычитатель и соответствующее управление, которое обеспечило бы вычитание меньшего числа из большего. Накопление упрощается при использовании вместо прямого кода дополнительного.

Отсчеты и коэффициенты также могут храниться в дополнительном, а не в прямом коде. Это означает, что необходимо специальное управление первым сумматором (рис. 4.3). Удобнее всего для этого использовать алгоритм Бута, в котором анализ двух соседних разрядов позволяет решить, что следует выполнить – сложение или вычитание.

4.2. Параллелизм при построении КИХ-фильтров прямой формы

Если входные отсчеты представлены b -разрядными числами, для вычисления выходного отсчета приходится выполнить N умножений и предполагается, что время сдвига регистра на один разряд τ равно времени сложения, то максимальная скорость обработки данных R , которая может быть достигнута с помощью схемы, рассмотренной в 4.1 (см., например, рис. 4.3), равна $1/(Nbt)$. Так, если $\tau = 100$ нс, $N = 16$ и $b = 16$, то

$$R = \frac{10^9}{256 \times 100} = 39062,5 \text{ Гц}.$$

Для увеличения R можно использовать различные способы параллелизма, касающиеся как работы памяти, так и выполнения арифметических операций. Используя два или более регистра сдвига, можно пропорционально увеличить скорость обработки; аналогичного результата можно достичь, используя два или более арифметических устройства (АУ). Рассмотрим все эти варианты, чтобы выяснить, какое дополнительное управление необходимо и какими должны быть арифметические устройства.

На рис. 4.4 представлены все необходимые элементы и связи между ними для системы, содержащей два параллельно работающих арифметических устройства. Там же показана последовательность вычислений и размещение данных в регистре сдвига после выполнения очередной операции умножение–накопление.

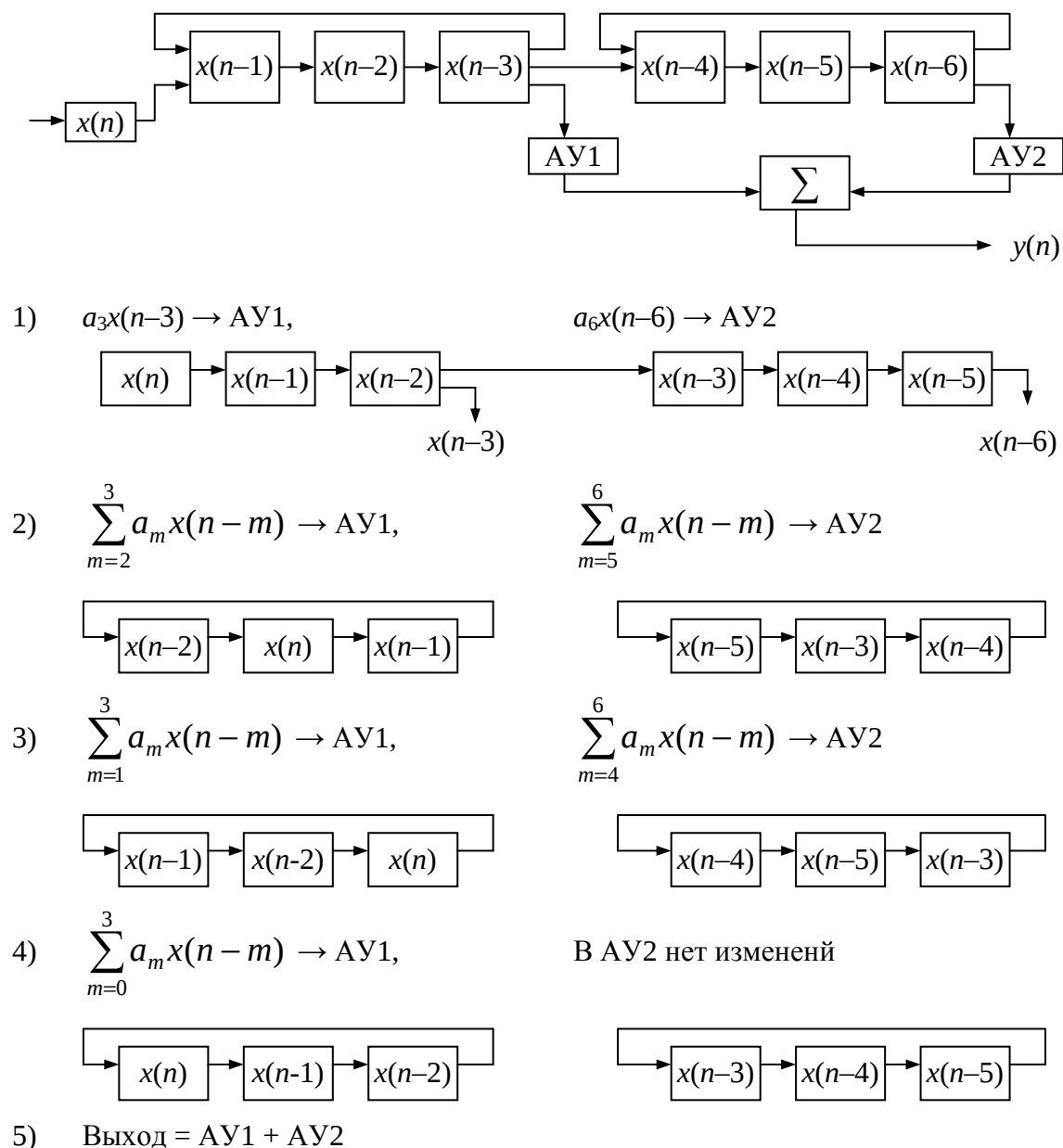


Рис. 4.4. – Структурная схема, последовательность обработки и промежуточные результаты для КИХ-фильтра с двумя параллельно работающими АУ

По сравнению со схемой, содержащей одно АУ (она была приведена на рис. 4.3), данная схема обеспечивает вдвое большую скорость обработки при условии, что схемы АУ1 и АУ2 такие же, как у арифметического устройства на рис. 4.3. Рассмотренный метод параллелизма легко обобщить на любое число арифметических устройств, однако не следует забывать, что высокий уровень интеграции сохранится только в том случае, если каждая из частей регистра сдвига является достаточно многоразрядной. Так, при построении фильтра 20-го порядка с 16-разрядными отсчетами и одним АУ требуется 320-разрядный кристалл. При построении такого же фильтра с двумя АУ необходимо иметь регистр сдвига с отводом, что при современной номенклатуре микросхем потребует использования двух корпусов, хотя не исключено, что такие регистры сдвига с отводами вполне могут быть изготовлены специально.

4.3. Каскадная форма КИХ-фильтра

Схема, приведенная на рис. 4.4, может быть использована и для построения КИХ-фильтров каскадной формы. На рис. 4.5 в качестве примера изобра-

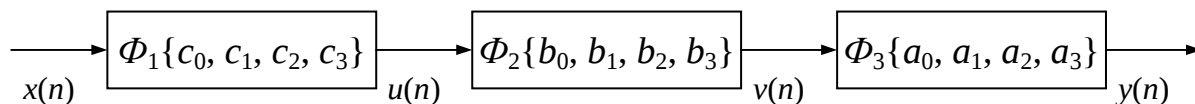


Рис. 4.5. – Трехкаскадная форма КИХ-фильтра

жен трехкаскадный фильтр, причем каждый из блоков представляет собой фильтр четвертого порядка. На рис. 4.6 показано, как строится такой фильтр с использованием одного АУ и одного последовательного регистра сдвига в качестве памяти промежуточных результатов. Здесь требуется сохранять выходные отсчеты каждого из блоков и вводить их в соответствующие моменты времени в регистр сдвига. Характер размещения промежуточных результатов в регистре сдвига для 12 последовательных тактов обработки показан на рис. 4.7а, а программа, по которой работают все устройства фильтра, – на рис. 4.7б. Видно, что отсчеты $x(n-4)$, $u(n-4)$ и $v(n-4)$ выводятся из памяти, а вместо них вводятся три новых отсчета.

Фильтр, изображенный на рис. 4.5, легко может быть построен с использованием параллельно работающих арифметических устройств, причем здесь возможны два простых подхода. Первый из них состоит в том, что для каждого из блоков четвертого порядка используется свое АУ, так что весь фильтр будет содержать три АУ, как показано на рис. 4.8. Из приведенной там же

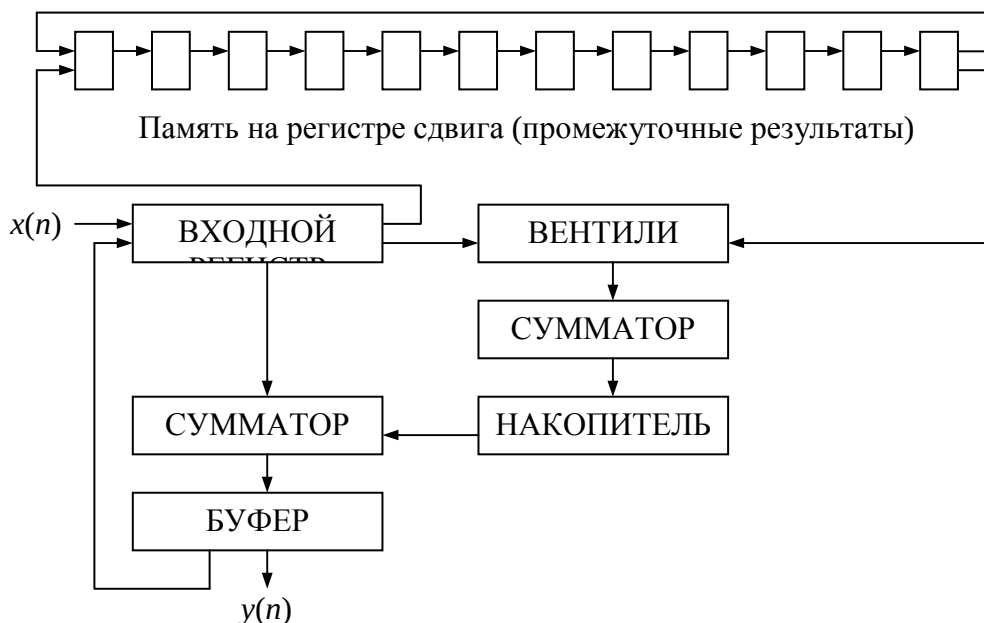


Рис. 4.6. – Построение каскадной формы КИХ-фильтра с использованием единственного АУ

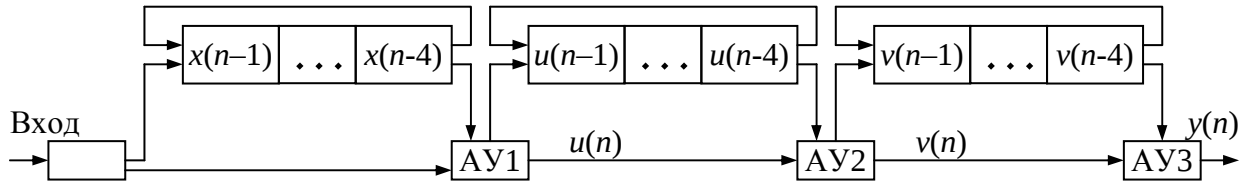
	$\overline{v(n-1)}$	$\overline{v(n-1)}$	$\overline{v(n-2)}$	$\overline{v(n-3)}$	$\overline{v(n-4)}$	$\overline{u(n-1)}$	$\overline{u(n-2)}$	$\overline{u(n-3)}$	$\overline{u(n-4)}$	$\overline{x(n-1)}$	$\overline{x(n-2)}$	$\overline{x(n-3)}$
1) —	$v(n-1)$	$v(n-2)$	$v(n-3)$	$v(n-4)$	$u(n-1)$	$u(n-2)$	$u(n-3)$	$u(n-4)$	$x(n-1)$	$x(n-2)$	$x(n-3)$	$x(n-4)$
2) $x(n-3)$	—	$v(n-1)$	$v(n-2)$	$v(n-3)$	$v(n-4)$	$u(n-1)$	$u(n-2)$	$u(n-3)$	$u(n-4)$	$x(n-1)$	$x(n-2)$	$x(n-3)$
3) $x(n-3)$	$x(n-3)$	—	$v(n-1)$	$v(n-2)$	$v(n-3)$	$v(n-4)$	$u(n-1)$	$u(n-2)$	$u(n-3)$	$u(n-4)$	$x(n-1)$	$x(n-2)$
4) $x(n-2)$	$x(n-3)$	$x(n-3)$	—	$v(n-1)$	$v(n-2)$	$v(n-3)$	$v(n-4)$	$u(n-1)$	$u(n-2)$	$u(n-3)$	$u(n-4)$	$x(n-1)$
5) $x(n-1)$	$x(n-2)$	$x(n-3)$	$x(n-3)$	—	$v(n-1)$	$v(n-2)$	$v(n-3)$	$v(n-4)$	$u(n-1)$	$u(n-2)$	$u(n-3)$	$u(n-4)$
6) $u(n-3)$	$x(n-1)$	$x(n-2)$	$x(n-3)$	$x(n-3)$	—	$v(n-1)$	$v(n-2)$	$v(n-3)$	$v(n-4)$	$u(n-1)$	$u(n-2)$	$u(n-3)$
7) $u(n-3)$	$u(n-3)$	$x(n-1)$	$x(n-2)$	$x(n-3)$	$x(n-3)$	—	$v(n-1)$	$v(n-2)$	$v(n-3)$	$v(n-4)$	$u(n-1)$	$u(n-2)$
8) $u(n-2)$	$u(n-3)$	$u(n-3)$	$x(n-1)$	$x(n-2)$	$x(n-3)$	$x(n-3)$	$x(n-3)$	—	$v(n-1)$	$v(n-2)$	$v(n-3)$	$v(n-4)$
9) $u(n-1)$	$u(n-2)$	$u(n-3)$	$u(n-3)$	$x(n-1)$	$x(n-2)$	$x(n-3)$	$x(n-3)$	$x(n-3)$	—	$v(n-1)$	$v(n-2)$	$v(n-3)$
10) $v(n-3)$	$u(n-1)$	$u(n-2)$	$u(n-3)$	$u(n-3)$	$u(n-3)$	$x(n-1)$	$x(n-2)$	$x(n-3)$	$x(n-3)$	—	$v(n-1)$	$v(n-2)$
11) $v(n-3)$	$v(n-3)$	$u(n-1)$	$u(n-2)$	$u(n-3)$	$u(n-3)$	$x(n-1)$	$x(n-2)$	$x(n-3)$	$x(n-3)$	$x(n-3)$	—	$v(n-1)$
12) $v(n-2)$	$v(n-3)$	$v(n-3)$	$v(n-3)$	$u(n-1)$	$u(n-2)$	$u(n-3)$	$u(n-3)$	$x(n-1)$	$x(n-2)$	$x(n-3)$	$x(n-3)$	—
13) $x(n)$	$v(n-1)$	$v(n-2)$	$v(n-3)$	$v(n-3)$	$v(n-3)$	$u(n-1)$	$u(n-2)$	$u(n-3)$	$u(n-3)$	$x(n-1)$	$x(n-2)$	$x(n-3)$

Рис. 4.7а. – Последовательные состояния памяти для каскадной формы КИХ-фильтра

программы работы фильтра видно, что во всех трех АУ расчет последовательных частичных сумм произведений производится до тех пор, пока в каждом из них не будет сформирована сумма всех четырех произведений. На пятом такте в первом АУ заканчивается вычисление $u(n)$, на вход поступает новый отсчет $x(n)$ и все регистры сдвигаются. Получаемое значение $u(n)$ используется в качестве входного для следующего блока и может быть добавлено к результату второго АУ и т. д. Таким образом, при использовании трех последовательно включенных арифметических устройств для завершения вычислений необходимы три дополнительных такта.

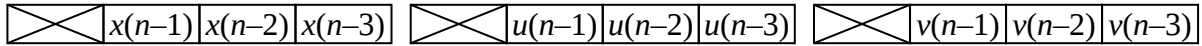
- 1) $c_4x(n-4)$ → Накопитель
- 2) $c_4x(n-4) + c_3x(n-3)$ → Накопитель
- 3) $\sum_{m=2}^4 c_m x(n-m)$ → Накопитель
- 4) $\sum_{m=1}^4 c_m x(n-m)$ → Накопитель
- 5) $u(n) = \sum_{m=0}^4 c_m x(n-m)$; Ввести $x(n)$ в ЗУ; Ввести $u(n)$ во входной регистр;
 $b_4u(n-4)$ → Накопитель
- 6) $b_4u(n-4) + b_3u(n-3)$ → Накопитель
- 7) $\sum_{m=2}^4 b_m u(n-m)$ → Накопитель
- 8) $\sum_{m=1}^4 b_m u(n-m)$ → Накопитель
- 9) $v(n) = \sum_{m=0}^4 b_m u(n-m)$; Ввести $u(n)$ в ЗУ; Ввести $v(n)$ во входной регистр;
 $a_4v(n-4)$ → Накопитель
- 10) $a_4v(n-4) + a_3v(n-3)$ → Накопитель
- 11) $\sum_{m=2}^4 a_m v(n-m)$ → Накопитель
- 12) $\sum_{m=1}^4 a_m v(n-m)$ → Накопитель
- 13) $y(n) = \sum_{m=0}^4 a_m v(n-m)$; Выход $y(n)$;
Ввести $v(n)$ в ЗУ; Ввести $x(n)$ во входной регистр

Рис. 4.76. – Программа работы каскадной формы КИХ-фильтра

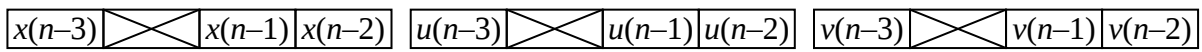


Программа

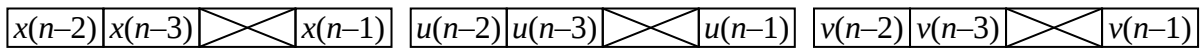
- 1) Вычислить $x(n-4)c_4$, $u(n-4)b_4$, $v(n-4)a_4$ и вытолкнуть из регистров $x(n-4)$, $u(n-4)$, $v(n-4)$



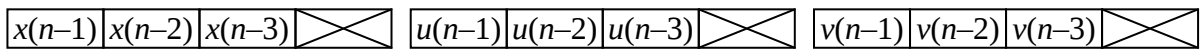
- 2) Вычислить $\sum_{i=3}^4 c_i x(n-i)$, $\sum_{i=3}^4 b_i u(n-i)$, $\sum_{i=3}^4 a_i v(n-i)$ и осуществить сдвиг во всех блоках



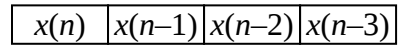
- 3) Вычислить $\sum_{i=2}^4 c_i x(n-i)$, $\sum_{i=2}^4 b_i u(n-i)$, $\sum_{i=2}^4 a_i v(n-i)$ и осуществить сдвиг во всех блоках



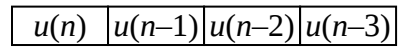
- 4) Вычислить $\sum_{i=1}^4 c_i x(n-i)$, $\sum_{i=1}^4 b_i u(n-i)$, $\sum_{i=1}^4 a_i v(n-i)$ и осуществить сдвиг во всех блоках



- 5) Вычислить $u(n) = \sum_{i=0}^4 c_i x(n-i)$ и ввести $x(n)$ в первый блок



- 6) Вычислить $v(n) = \sum_{i=0}^4 b_i u(n-i)$ и ввести $u(n)$ в второй блок



- 7) Вычислить $y(n) = \sum_{i=0}^4 a_i v(n-i)$ и ввести $v(n)$ в третий блок

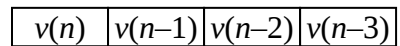


Рис. 4.8. – Построение каскадного КИХ-фильтра с использованием трех АУ

4.4. Прямая форма КИХ-фильтра с высоким уровнем параллелизма

Как уже было отмечено, основная структура фильтра с регистром сдвига в качестве ЗУ позволяет использовать параллелизм и при выполнении арифметических операций, и при работе памяти. Возникает вопрос, какого предельного уровня параллелизма можно достичь в схемах подобного типа. Прежде чем ответить на этот вопрос, рассмотрим систему с достаточно очевидным параллелизмом, а именно систему N -го порядка с N арифметическими устройствами. В такой системе, изображенной на рис. 4.9, для каждого из входных отсчетов используется отдельный умножитель. Заметим, что в этом

случае не требуется кругового сдвига памяти, так как регистр используется лишь как обычная линия задержки с отводами. Отметим также, что необходимость иметь отводы сводит на нет преимущества современной БИС-технологии. Как и во всех предыдущих примерах, отсчеты в регистре сдвига

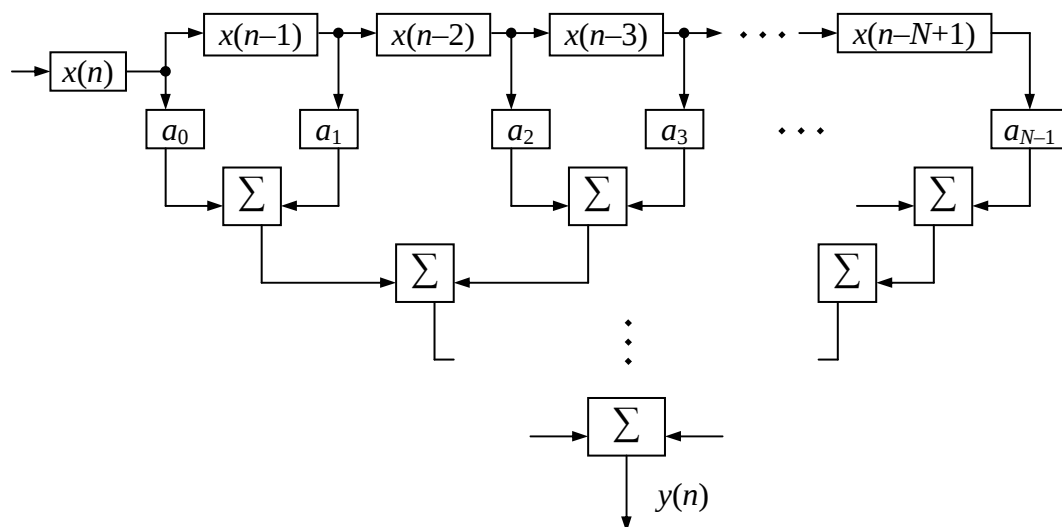


Рис. 4.9. – Структурная схема КИХ-фильтра N -го порядка с N параллельно работающими АУ

могут храниться как в последовательном, так и в параллельном виде.

Далее, кроме N умножителей, необходимо иметь столько же сумматоров. Считая, что время умножения согласовано со временем сдвига регистра из одного состояния в другое, приходим к выводу, что необходимо иметь дополнительное время для накопления произведений. Это дополнительное время может не потребоваться, если в точках схемы, отмеченных на рис. 4.9, ввести буферные регистры, что приводит к организации обработки в поточной форме. Если предположить, что общее время накопления в точности равно времени умножения, то выходная последовательность будет дополнительно задержана на один отсчет. Приведенная на рис. 4.9 схема – не единственная, позволяющая для построения параллельного КИХ-фильтра N -го порядка использовать N арифметических устройств. Другая возможная схема показана на рис. 4.10.

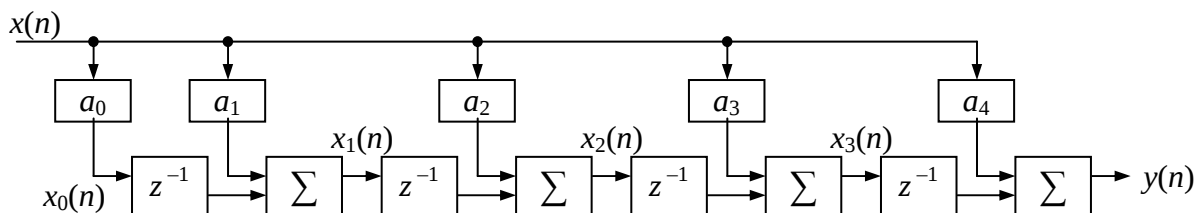


Рис. 4.10. – Вариант построения КИХ-фильтра

Эта схема описывается следующими уравнениями

$$x_1(n) = a_0 x(n-1) + a_1 x(n),$$

$$x_2(n) = x_1(n-1) + a_2 x(n) = a_0 x(n-2) + a_1 x(n-1) + a_2 x(n),$$

$$x_3(n) = x_2(n-1) + a_3 x(n) = \sum_{i=0}^3 a_i x(n-3+i),$$

$$y(n) = x_3(n-1) + a_4 x(n) = \sum_{i=0}^4 a_i x(n-4+i).$$

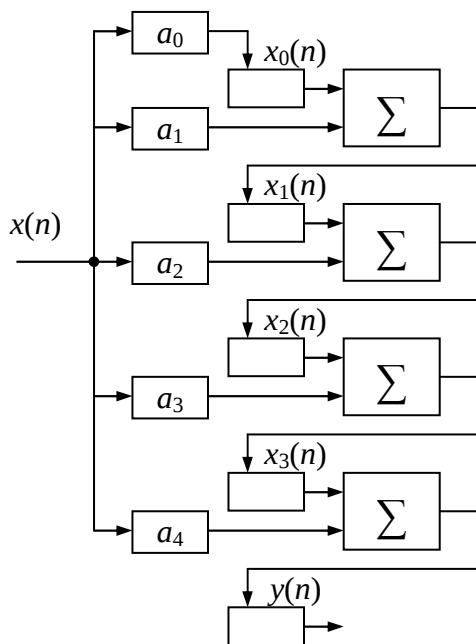


Рис. 4.11. – Структурная схема КИХ-фильтра

Видно, что, как и раньше, результат равен свертке, хотя индексы и оказались переставленными. К обычной форме можно перейти при $b_0 = a_3$, $b_1 = a_2$, $b_2 = a_1$, $b_3 = a_0$. На рис. 4.11 показана простая схема построения фильтра, приведенного на рис. 4.10.

Из схемы на рис. 4.11 вытекает простой способ аппаратного построения фильтра при условии, что отсчеты $x(n)$ поступают на умножители в последовательной форме и поразрядно умножаются на коэффициенты. В этом случае произведения могут формироваться последовательно, что позволяет применить последовательные сумматоры. Таким образом, по мере того как отсчеты $x(n)$ поразрядно выталкиваются из регистра в сумматор, в этот же регистр в последовательной форме вводится результат суммирования значения $x_{i-1}(n)$ с произведением $a_i x(n)$.

4.5. Прямая форма построения БИХ-фильтров

До сих пор рассматривались структуры, соответствующие КИХ-фильтрам. Для перехода к БИХ-фильтрам в эти структуры необходимо ввести лишь незначительные изменения. В качестве примера рассмотрим построение фильтра, содержащего только полюсы, используя форму БИХ-фильтра, приведенную на рис. 4.12.

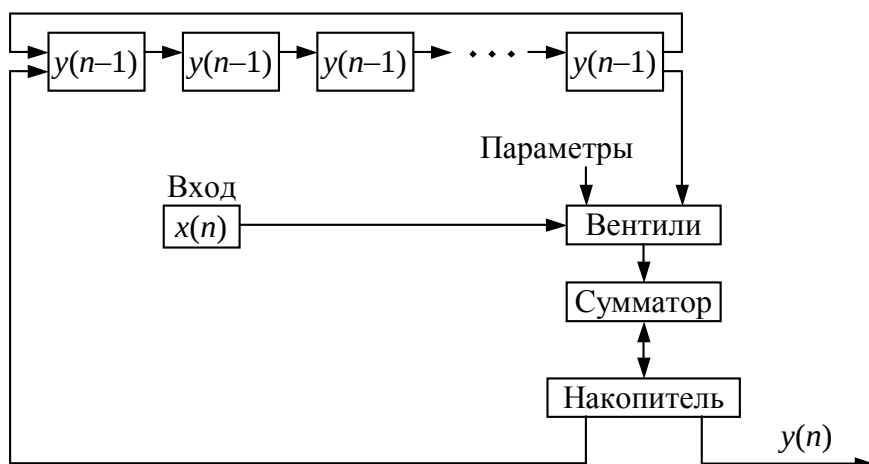


Рис. 4.12. – Прямая форма построения БИХ-фильтра

Считается, что коэффициенты используются при вычислениях в параллельной форме, тогда как промежуточные результаты хранятся в памяти, как и в случае КИХ-фильтров, в последовательной форме. Разностное уравнение изображенного на рис. 4.12 БИХ-фильтра, содержащего только полюсы, имеет вид

$$y(n) = \sum_{j=1}^{12} y(n-j)a_j + bx(n),$$

причем коэффициенты a_j и b считаются известными.

Промежуточные результаты $y(n-j)$ и входные отсчеты $x(n)$ поступают на входы в последовательной форме. При умножении $y(n-12)$ на a_{12} значение $y(n-12)$ поразрядно выдается с выхода регистра сдвига. При выполнении оставшихся 11 умножений, начиная с $y(n-11)a_{11}$ и кончая $y(n-1)a_1$, содержимое регистра сдвигается по кругу. Таким образом, перед поступлением на вход регистра $x(n)$ его состояние имеет вид, представленный на рис. 4.13.

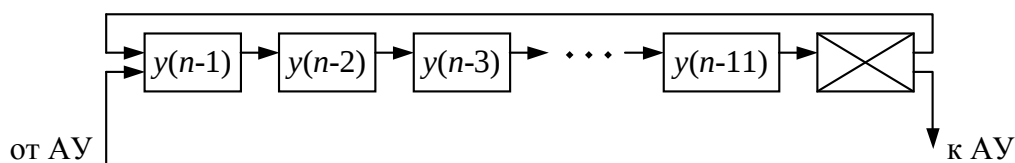


Рис. 4.13. – Состояние памяти БИХ-фильтра непосредственно перед вводом входного отсчёта $x(n)$

Пока вычисляется произведение $bx(n)$, отсчет $y(n)$ на вход регистра не подается, поэтому круговой сдвиг регистра не производится. Заключительный цикл включает ввод накопленного значения $y(n)$ в регистр сдвига, ввод нового значения $x(n)$, а также очистку накопителя непосредственно перед началом следующего цикла.

Примем для примера значение частоты дискретизации равным 8 кГц, что соответствует интервалу между отсчетами, равному 125 мкс. Если считать, что промежуточные результаты представляются 16-разрядными числами, то общее количество сдвигов регистра на один разряд оказывается равным $16 \times 14 = 224$, т. е. период следования тактовых импульсов должен быть равен $125/224$ мкс (0,55 мкс). Итак, информация в регистре сдвига должна циркулировать с частотой 2 Мбит/с; с такой же частотой должно производиться сложение в сумматоре. Такая частота вполне достижима при использовании микросхем с умеренным быстродействием, так что подобные фильтры вполне можно разработать в виде большой интегральной схемы.

4.6. Каскадная форма БИХ-фильтров

На рис. 4.14 изображен каскадный БИХ-фильтр, построенный из трех последовательно соединенных блоков второго порядка, каждый из которых имеет по два полюса и по два нуля. Все блоки второго порядка построены на основе прямой канонической формы, т. е. с использованием минимума элементов задержки.

Простая схема построения такого фильтра с использованием двух арифметических устройств, а также последовательность выполнения вычислений (т. е. управление) представлены на рис. 4.15. Эта схема аналогична рассмотренным выше в том смысле, что, как и ранее, память строится на последовательных регистрах сдвига, а умножитель использует сложения и сдвиги, обрабатывая последовательно поступающие разряды промежуточных результатов.

Удобство системы из двух АУ заключается в независимом вычислении

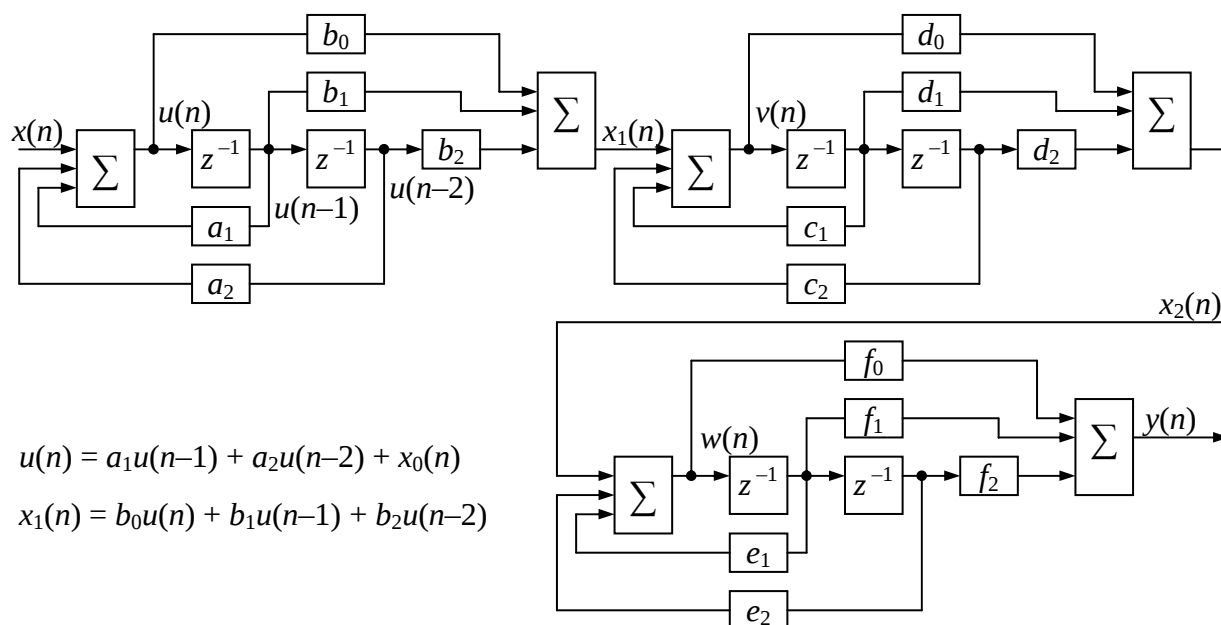


Рис. 4.14. – Каскадная форма БИХ-фильтра, состоящая из трех последовательно соединенных блоков второго порядка

результатов в обоих АУ последовательно выполняются две вспомогательные операции: одна для формирования недостающего отсчета для цепи прямой связи, а другая для формирования входного отсчета. следующего каскада.

В приведенной на рис. 4.15 последовательности вычислений отмечено, когда регистры сдвигаются и когда сдвигов не должно быть. Видно, что для фильтра шестого порядка в общей сложности требуется семь сдвигов, а также шесть интервалов времени, на которых сдвиги не производятся. Масштабирование (т. е. простой сдвиг отсчетов) $x_0(n)$, $x_1(n)$ и $x_2(n)$ также может потребоваться, когда эти отсчеты вводятся в процессе вычислений в АУ1.

Следует отметить, что в настоящее время алгоритмы обработки данных, чаще реализуются программным способом с использованием сигнальных процессоров.

5. ЦИФРОВОЕ СЖАТИЕ ВИДЕОСИГНАЛОВ

В настоящем разделе будут рассмотрены основные операции цифрового сжатия видеосигналов. Эти сведения являются введением в описание стандарта *MPEG*, приведенного в п. 5.9 [13].

5.1. Пространственная и временная избыточность

Для сжатия видеосигналов требуется идентифицировать избыточность типичного материала источника. Имеются две основные разновидности избыточности, которыми можно воспользоваться. Первая из них – это внутрикадровая, или пространственная избыточность. Ее можно идентифицировать по одному текущему видеокадру без обращения к какому-либо другому видеокадру. Вторая разновидность – межкадровая, или временная избыточность, для идентификации которой требуется как текущий, так и следующий видеокادر.

5.2. Дискретное косинусное преобразование

В каждом реальном телевизионном материале присутствует пространственная избыточность. Если изображение содержит объект достаточно большого размера, то все представляющие этот объект элементы имеют весьма близкие значения. Крупные объекты создают низкие пространственные частоты, а небольшие объекты – высокие пространственные частоты. Как правило, уровни всех этих частот не могут одновременно достигать больших значений. Обычный ИКМ-видеосигнал должен передавать любые пространственные частоты во всем диапазоне, однако если выполнить частотный анализ, то станет возможным передавать только фактически присутствующие частоты. Следовательно, при внутрикадровом кодировании одной из важных операций является выполнение анализа пространственных частот изображения.

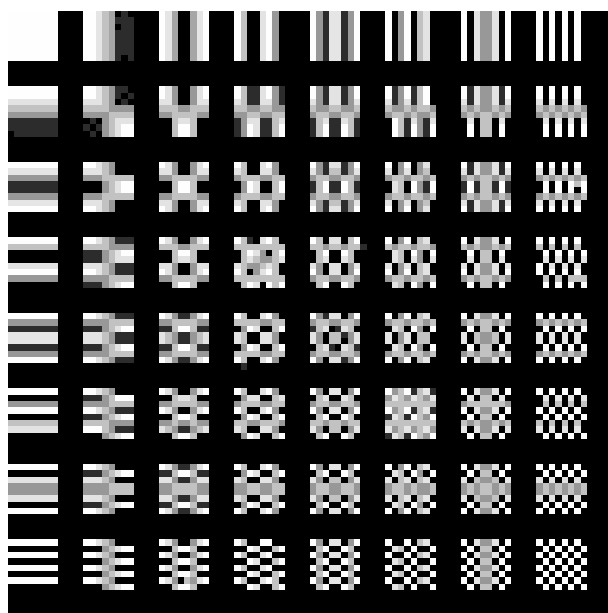


Рис. 5.1. – Двумерное дискретное косинусное преобразование

Следовательно, при внутрикадровом кодировании одной из важных операций является выполнение анализа пространственных частот изображения.

В стандарте *MPEG* применено дискретное косинусное преобразование (ДКП). Рис. 5.1 иллюстрирует действие двумерного ДКП. Изображение преобразуется блок за

блоком. Типичный блок содержит 8×8 отсчетов изображения. В результате ДКП такого блока образуется блок из 64 коэффициентов. Коэффициент – это число, выражающее степень присутствия конкретной пространственной частоты, имеющейся в изображении. На рисунке показаны блоки элементов изображения, соответствующие каждому из коэффициентов. Верхний левый коэффициент характеризует среднюю яркость блока и является величиной, пропорциональной арифметическому среднему значений всех элементов изображения, то есть пропорциональной постоянной составляющей. Коэффициенты, идущие далее слева направо, представляют повышающиеся горизонтальные пространственные частоты. Коэффициенты, идущие вниз, представляют повышающиеся вертикальные пространственные частоты.

Очевидно, что ДКП само по себе никакого сжатия не дает. Более того, длины кодовых слов для коэффициентов больше, чем для исходных отсчетов изображения. Однако ДКП обеспечивает преобразование отсчетов изображения в форму, позволяющую идентифицировать избыточность. Поскольку не все пространственные частоты присутствуют одновременно, то ДКП формирует набор коэффициентов, часть из которых имеет сравнительно большие значения, однако многие коэффициенты малы или равны нулю.

Очевидно, что если коэффициент равен нулю, то неважно, передается он или нет. Если коэффициент имеет малое (почти нулевое) значение, то его исключение дает тот же эффект, как и добавление слабой помехи с той же пространственной частотой, но с противоположным знаком. Решение исключить какой-то коэффициент принимается на основе видности такого нежелательного сигнала малой амплитуды. Если коэффициент очень большой и не может быть исключен, сжатие все же может быть обеспечено благодаря уменьшению числа битов, используемых для передачи этого коэффициента. Эта операция также сопровождается определенным негативным эффектом, а именно к изображению добавляется шум небольшого уровня.

5.3. Взвешивание

Видность пространственных частот, наблюдаемых человеком, неодинакова; на высоких частотах допустимы намного более высокие по величине уровни шума. Поэтому при кодировании применяется операция взвешивания, обеспечивающая концентрацию любого возникающего шума в области высоких частот. Каждый коэффициент делится на вес, который является функцией его позиции в блоке. Постоянная составляющая вообще не взвешивается, а вес коэффициента (делитель) возрастает по мере приближения к нижнему правому углу. В декодере необходимо произвести операцию, обратную взвешиванию. При этом коэффициенты для более высоких частот умножаются на большие числа, что приводит к повышению шума на этих высоких частотах.

После взвешивания некоторые небольшие коэффициенты становятся еще меньше. В типичном телевизионном материале коэффициенты с большими

значениями чаще всего находятся в верхнем левом углу, а остальные коэффициенты часто пренебрежимо малы или равны нулю. Поэтому выгодно передавать коэффициенты в зигзагообразной последовательности, начиная с верхнего левого угла. В результате этого процесса ненулевые коэффициенты, как правило, передаются первыми, а нулевые – в самом конце. В какой-то точке, начиная с которой все последующие коэффициенты равны нулю, имеет смысл вообще прервать передачу с помощью простого символа, сообщающего приемнику, что ненулевых коэффициентов дальше не будет.

5.4. Кодирование с переменной длиной слова

Анализ типичных исходных данных показывает, что вероятность появления коэффициентов с малыми значениями, имеющих много нулей в начале слов, выше вероятности появления коэффициентов с большими значениями, передаваемыми согласно правилам, описанным в разделе 5.2. Когда вероятность появления таких значений известна, можно воспользоваться способом кодирования, который извлекает пользу из этого знания распределения вероятностей.

Кодирование по Хаффмену предусматривает назначение коротких кодовых слов более вероятным значениям и длинных кодовых слов значениям, которые менее вероятны. Если исходные данные соответствуют предполагаемому распределению вероятностей, то этот метод обеспечивает повышение эффективности передачи.

При кодировании с ДКП часто случается так, что многие коэффициенты имеют нулевое значение. В этом случае кодеру длин серий остается только сообщить декодеру количество нулевых коэффициентов вместо того, чтобы передавать их по отдельности.

5.5. Внутрикадровое кодирование

На рис. 5.2 показана полная структурная схема внутрикадрового кодирования. Входное изображение в виде строк растра преобразуется в совокуп-

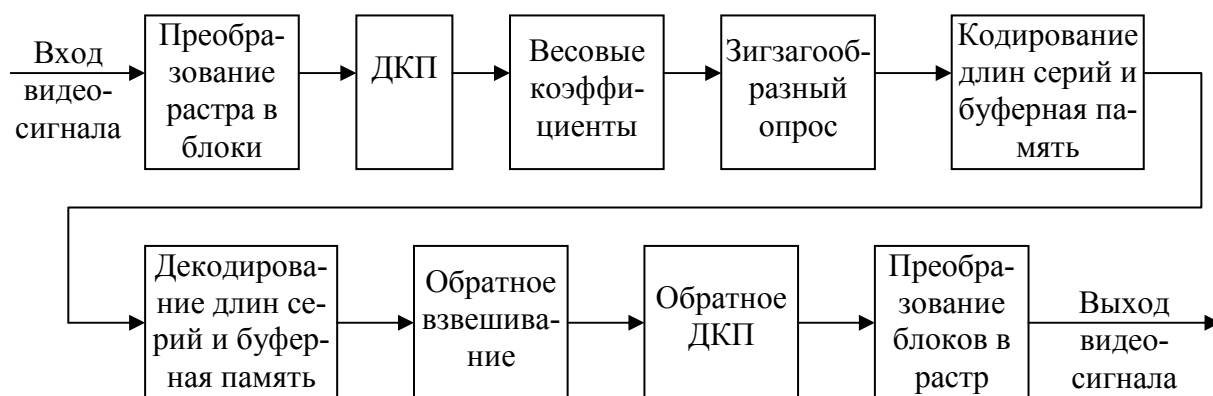


Рис. 5.2. – Полная структурная схема внутрикадрового кодера

ность блоков. Блоки подвергаются ДКП. Затем полученные коэффициенты опрашиваются зигзагообразно, взвешиваются перед их переквантованием (укорочением кодовых слов) и подвергаются кодированию длин серий. В практических системах требуется иметь постоянный выходной цифровой поток, если даже энтропия источника варьируется.

Возможны два решения. Первое из них – применить буферное запоминающее устройство, «поглощающее» такие изменения. Другое решение состоит в измерении объема выходных данных и сравнении результата с требуемым объемом данных. Если данных слишком много, применяют переквантование для более значительного уменьшения длины слова коэффициентов. Если данных недостаточно, переквантование делают менее эффективным.

На рис. 5.2 также показана структурная схема соответствующего декодера. Сначала декодируются данные, подвергнутые кодированию длин серий, а затем производится компенсация взвешивания коэффициентов и их группирование в блоки. В результате обратного преобразования формируется один блок элементов изображения, который записывается в запоминающее устройство (ЗУ) вместе с другими блоками, что обеспечивает восстановление растровой структуры выходного изображения.

5.6. Межкадровое кодирование

Система сжатия на рис. 5.2, в основном, не отличается от системы, разработанной группой *JPEG* для неподвижных изображений. Внутрикадровое сжатие осуществляется без использования избыточности, относящейся к смежным видеокадрам. Однако в типичном видеоматериале текущий видеокادر часто очень похож на предыдущий видеокادر. Рис. 5.3 иллюстрирует работу дифференциального кодера. Для задержки предыдущего видеокадра необходимо предусмотреть запоминающее устройство на кадр. Каждый отсчет предыдущего видеокадра вычитается из соответствующих отсчетов текущего видеокадра, в результате чего формируется разностный видеокادر. Разностный видеокادر является полноправным изображением, и его можно подвергнуть процедуре сжатия с использованием внутрикадрового кодирования на основе ДКП. В декодере имеется другое кадровое ЗУ, которое задерживает предыдущий выходной видеокادر. Декодированный разностный видеокادر добавляется к предыдущему видеокадру, чтобы сформировать теку-

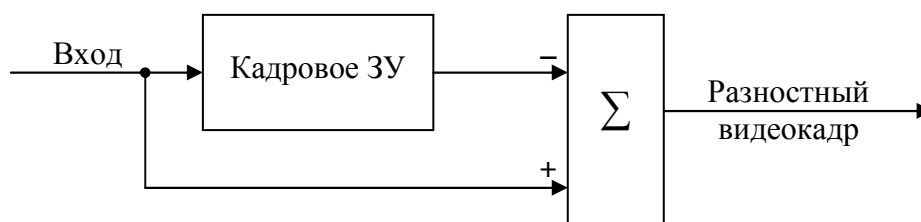
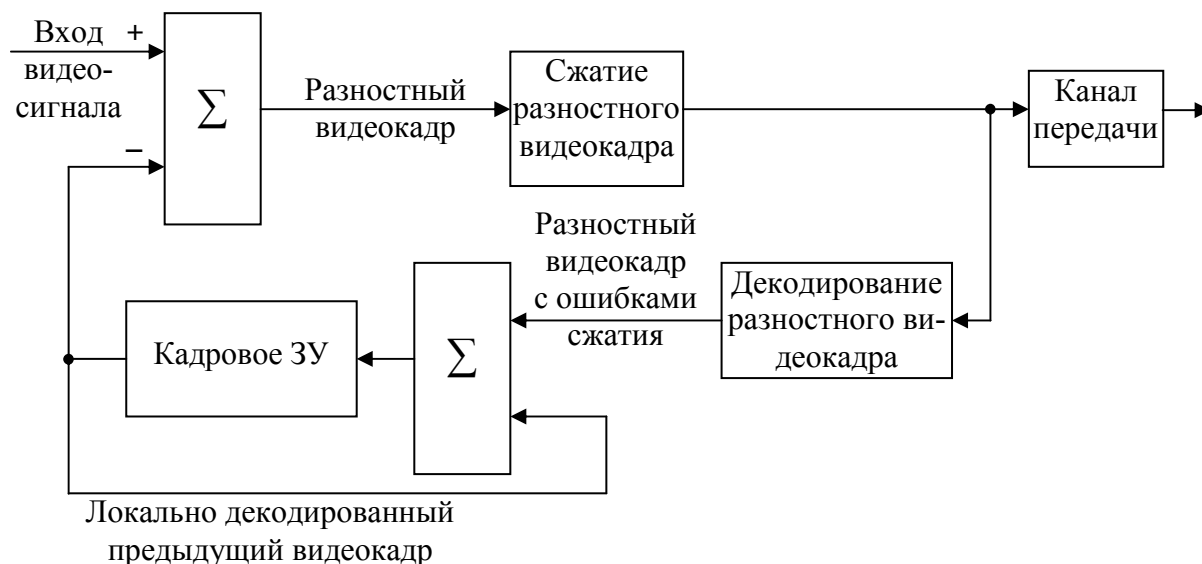


Рис. 5.3. – Дифференциальный кодер

щий видеокадр, который не только подается на выход, но и поступает в кадровое ЗУ, становясь основой формирования следующего видеокадра.

При разностном кодировании важно предотвратить рост небольших ошибок, которые могут накапливаться в серии видеокадров. Одним из источников ошибок является процесс сжатия разностного видеокадра. Рис. 5.4 иллюстриру-

а)



б)

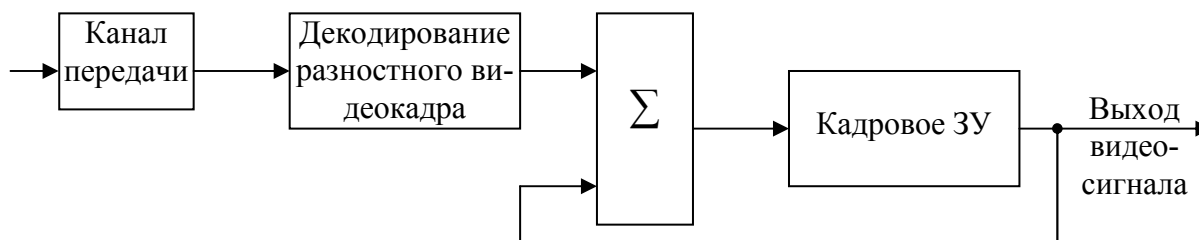


Рис. 5.4. – Кодер (а) и декодер (б) сжатия видеоданных

ет способ устранения таких ошибок. Кодер и декодер сжатия содержат идентичные декодеры разностного видеокадра. Когда в кодере производится вычитание предыдущего видеокадра из текущего, то фактически это означает вычитание локально декодированного сжатого видеокадра, идентичного тому, которое будет иметь реальный декодер. В результате любые небольшие ошибки, вызываемые сжатием предыдущих разностных видеокадров, устраняются.

5.7. Компенсация движения

Дифференциальное кодирование эффективно применительно к материалу, содержащему, в основном, неподвижные изображения, однако становится неэффективным при наличии движения. В предельном случае при панорамировании сцены изменяются значения буквально каждого отсчета изображе-

ния, поэтому разностные видеокadres содержат почти тот же объем данных, что и исходные. Следовательно, для достижения больших коэффициентов сжатия при дифференциальном кодировании необходимо применять компенсацию движения.

Как показано на рис. 5.5, в кодере сжатия с компенсацией движения сравниваются участки двух последовательных видеокadres, что позволяет определить направление и расстояние относительного движения между двумя кадрами. Эти данные представляются в виде двумерного вектора. Вектор движения используется в кодере для сдвига отсчетов предыдущего видеокadra и формирования так называемого предсказанного видеокadra, который затем вычитается из текущего, или фактического видеокadra. При точном измерении движения процедура сдвига обеспечит совмещение объектов в предсказанном видеокadre с теми же объектами в фактическом видеокadre, и благодаря этому объем разностных данных резко снижается. Очевидно, что если совместно с разностными данными передавать векторы движения, декодер сможет выполнить тот же самый сдвиг предыдущего видеокadra и сформировать локально предсказанный видеокadre. Для получения декодированного текущего кадра достаточно добавить разностные данные.

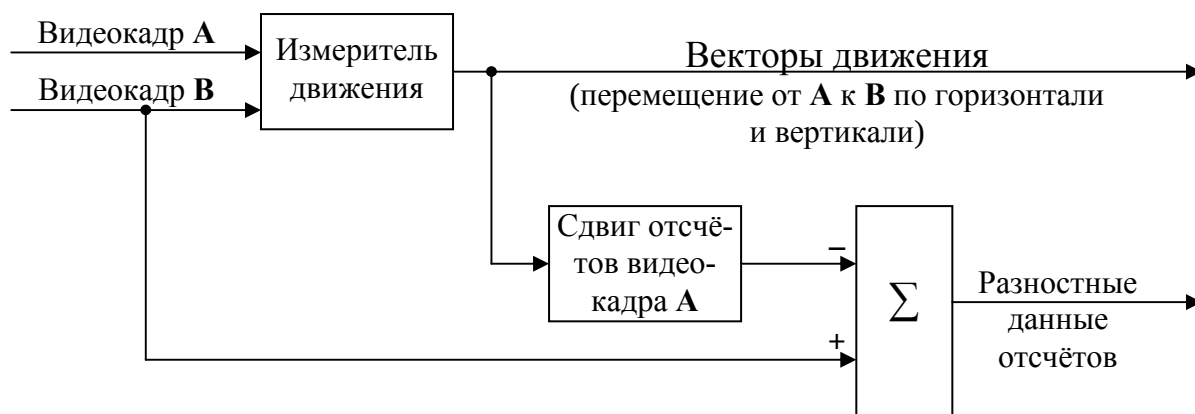


Рис. 5.5. – Компенсация движения в кодере сжатия

Видеокадр разбивается на прямоугольные участки, называемые макроблоками. Каждый такой макроблок снабжается индивидуальным вектором движения, воздействующим на весь блок целиком. Очевидно, что если в макроблок попадает край движущегося объекта, то одна его часть при движении изменяется, а другая — нет. Возможны два решения: а) можно обнулить векторы движения и истратить часть ресурса разностных данных на передачу движения; б) можно применить векторы движения к неподвижной части, что также увеличит объем разностных данных. «Умный» кодер сжатия может попытаться применить оба подхода, чтобы определить, какой из них обеспечивает меньший цифровой поток.

Поскольку в реальном видеоматериале движение имеет тенденцию быть относительно неизменным на интервале нескольких видеокadres, сами векторы движения тоже передаются разностным методом. Таким образом, нуле-

ное значение параметра вектора движения интерпретируется как продолжение равномерного движения. Векторы для предыдущего видеокадра просто копируются. Разностные векторы движения передаются только в случае изменения его скорости или направления.

5.8. I-видеокадры

Чисто дифференциальные кодеры сжатия непригодны для практического использования, поскольку декодер способен правильно интерпретировать передаваемые данные только в случае их приема с самого начала. Любые ошибки передачи будут затем влиять на изображение бесконечно долго. В практических дифференциальных системах передачи разностные видеокадры периодически перемежаются опорными («абсолютными») видеокадрами, подвергнутыми только внутрикадровому кодированию. Рис. 5.6 иллюстрирует этот принцип. Видеокадры с внутрикадровым кодированием, или *I*-видеокадры, формируются посредством выключения системы предсказания, и в этом случае предыдущие данные не вычитаются из входных данных. Разностный видеокадр становится опорным. Между *I*-видеокадрами действует механизм предсказания с компенсацией движения и формируются *P*-видеокадры, которые содержат векторы движения и значения разностных данных, сжатых методом ДКП.

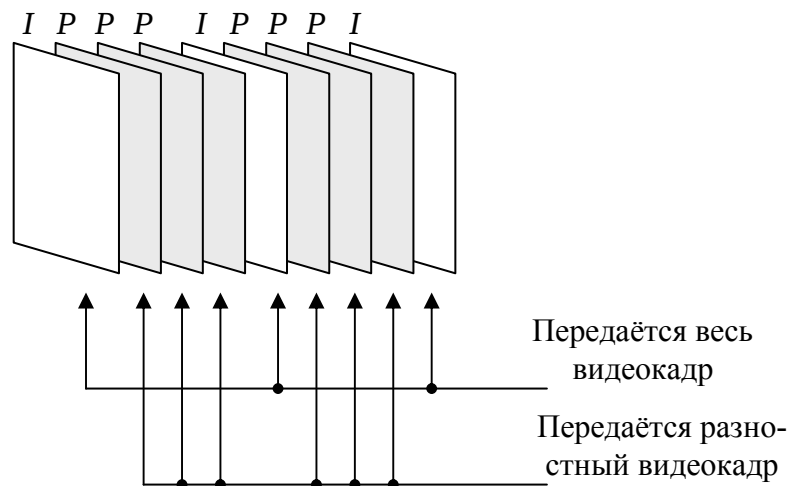


Рис. 5.6. – Передача опорных и разностных видеокадров

В случае видеокиноматериалов, создаваемых телекинодатчиком, кодеру сжатия полезно получить сведения о методе его развертки. Например, в телекинодатчике 60 Гц стандарта для получения частоты полей 60 Гц на основе частоты кадров 24 Гц используется протяжка фильма способом «3:2». При этом в каждом цикле из трех полей третье поле является прямым повторением первого поля, а следовательно, не несет новой информации. Кодеру сжатия достаточно передать только два первых поля. Последовательность «3:2» восстанавливается в декодере путем повторения одного из них.

В телекинодатчиках обоих стандартов (50 и 60 Гц) между двумя полями, источником которых является один и тот же кинокадр, движения быть не может. Следовательно, не имеет смысла измерять движение или передавать векторы движения для каждого второго поля.

Стандарт не требует передавать *I*-видеокадры с абсолютной регулярностью. Видеокадры разного типа маркируются индивидуально, поэтому декодер может автоматически воспринимать любой принятый видеокадр. Монтажные переходы в поступающем материале, выполненные методом быстрого переключения, вызывают появление разностных данных большого объема. Кодер сжатия, снабженный предпроцессором, способен предвидеть появление монтажного переключения в материале источника и обеспечить его временное совмещение с *I*-видеокадром; при этом достигается наиболее эффективное использование цифрового потока.

Из изложенного должно быть ясно, что монтаж цифрового потока, полученного дифференциальным кодированием, может оказаться проблематичным. Если переключение потока происходит в произвольной точке, декодер не может вырабатывать видеокадры до поступления очередного *I*-видеокадра. Следовательно, возможности видеомонтажа ограничиваются режимом переключения на *I*-видеокадрах. Этого явно недостаточно для целей видеопроизводства, поэтому в тех случаях, когда предполагается последующий монтаж, два поля чересстрочного кадра можно подвергнуть сжатию в виде одного *I*-видеокадра и одного *P*-видеокадра. При этом появляется возможность видеомонтажа с точностью до кадра. Ясно, что достижимый коэффициент сжатия будет относительно небольшим вследствие частого появления *I*-видеокадров.

5.9. Стандарты *MPEG*

Все идеи, представленные выше, сконцентрированы здесь в описании *MPEG*. *MPEG* – это не метод цифрового сжатия как таковой; более того, определить *MPEG* каким-то одним термином затруднительно. По существу, *MPEG* представляет собой набор стандартных средств, или точно определенных алгоритмов, которые могут комбинироваться многими способами при реализации аппаратуры цифрового сжатия.

5.9.1. Применения *MPEG*

Как ни странно, *MPEG* не определяет выбор методов сжатия для кодера. Вместо этого стандарты *MPEG* точно указывают декодеру, какой смысл приписывать тем или иным сжатым цифровым потокам. Стандарты *MPEG* не содержат указаний по передаче цифровых потоков, поскольку это определяется конкретным применением. Выбор такого подхода объясняется тем, что он предоставляет большую гибкость при сохранении совместимости. Промышленные фирмы могут пользоваться запатентованными алгоритмами кодирования и все же формировать стандартный цифровой поток. Алгоритмы могут

совершенствоваться все время, однако сигналы сохраняют совместимость с существующими декодерами.

Применение цифрового сжатия видеосигналов не ограничивается потребностями вещательного телевидения. Применения имеют широкий спектр, начиная с движущихся изображений малого размера в персональных компьютерах и кончая изображениями на больших экранах телевидения высокой четкости. Требования к качеству изображения могут быть самыми различными, начиная с относительно плохого качества для охранных и видеотелефонных целей с переходом к качеству интерактивных компакт-дисков (КД-И) и кончая «прозрачным» качеством. Иногда один и тот же передаваемый сигнал требуется декодировать на уровне различных стандартов. Например, нет смысла декодировать сигнал телевидения высокой четкости для отображения на экране телевизора с диагональю 31 см. От конкретного применения зависит и приемлемость объема инвестиций для обеспечения необходимой сложности кодера и декодера. Стандарты *MPEG* кроме цифрового сжатия видеосигналов охватывают звуковое сопровождение в сжатом виде и мультиплексирование аудио- и видеоданных таким образом, чтобы сохранялась синхронизация.

5.9.2. Профили и уровни

Требование иметь широкий диапазон качественных показателей и уровней сложности удовлетворяется благодаря подразделению стандартов *MPEG* на «профили» с последующим подразделением каждого профиля на «уровни». Как показывает табл. 5.1, профиль представляет метод, а уровень – некоторое ограничение этого метода, такое, как размер изображения или цифровой поток. Кроме того, в таблице указаны уровни, реализованные в настоящее время.

Таблица 5.1

		Уровни			
		Низкий, разрешение до 352 x 288 отсчетов до 4 Мбит/с	Основной, разрешение до 720 x 576 отсчетов до 15 Мбит/с	Высокий (1440 отсчетов), разрешение до 720 x 576 отсчетов до 60 Мбит/с	Высокий, разрешение до 1920 x 288 отсчетов до 80 Мбит/с
Профили	Простой				
	Основной		О.П.О.У		
	Адаптивный по шуму				
	Адаптивный по разрешению				
	Высокий				

Простой профиль, как указывает само название, дает возможность пользоваться кодером и декодером минимальной сложности. Системы с простым профилем, в основном, характеризуются тем, что в них используются только *I*- и *P*-видеокадры, рассмотренные выше.

Основной профиль разработан с целью быть пригодным для значительной части применений, он привлек к себе большое внимание и будет подробно описан далее.

Профиль, адаптивный к шуму, представляет собой систему, в которой для передачи одного видеосигнала используются два отдельных цифровых потока. Эти два потока можно декодировать совместно и получить высококачественное изображение и высокое отношение сигнал-шум. Однако в случае отсутствия приема одного из потоков результирующее изображение будет иметь ту же пространственную разрешающую способность, но отношение сигнал-шум окажется пониженным. Соответствующий кодер сжатия спроектирован таким образом, что он формирует не один, а два сжатых цифровых потока. Один из них является носителем зашумленного изображения, другой – сигнала «цифровой поддержки», который добавляется к первому сигналу для устранения шума. Ожидается, что этот метод найдет применение в ЦТВ (цифровом телевизионном вещании) с использованием передатчиков, излучающих сигнал с адаптацией к отношению сигнал-шум, первый цифровой поток которого передается с увеличенной мощностью. Стационарные приемники с хорошими антеннами смогут принимать оба цифровых потока для формирования качественного изображения. В случае плохого приема только более сильный сигнал достигнет выхода, позволяя получить изображение более низкого качества, но это лучше, чем не иметь изображения вовсе.

Профиль с адаптацией по разрешающей способности представляет систему, в которой предусматриваются несколько цифровых потоков, обеспечивающих возможность воспроизводить изображение с различной разрешающей способностью в соответствии с используемым устройством отображения. Один поток может декодироваться независимо для формирования изображения низкого разрешения, отображаемого на небольших экранах портативных телевизоров. Второй поток переносит сигналы «пространственной поддержки», которые при добавлении первого сигнала обеспечивают повышение разрешающей способности с целью отображения на больших экранах.

Высокий профиль адаптируется по частоте развертки, позволяя осуществить передачу цифрового потока, который можно декодировать с получением двух различных кадровых частот в соответствии со сложностью декодера.

Как показывает таблица, структура профиля позволяет иметь несколько уровней. Уровни различаются, в основном, реализуемой разрешающей способностью и требуемым цифровым потоком. Основной уровень пригоден для телевидения стандартной четкости (ТСЧ). Следовательно, большинство вещательных организаций и создателей программ заинтересовано именно в ва-

рианте О.П.О.У («основной профиль, основной уровень») подмножества стандартов *MPEG*.

В общем случае, декодер *MPEG* с заданными профилем и уровнем должен быть способен также декодировать входные сигналы, соответствующие более низким профилям и уровням. Поэтому адаптивный декодер будет нормально работать при получении сигнала основного или простого уровней.

В основе большинства вариантов иерархии *MPEG* лежит видеостандарт 4:2:0, предусматривающий вертикальную субдискретизацию цветоразностных сигналов. Широко распространилось мнение, что для целей ТВ вещания требуется работать по стандарту 4:2:2; поэтому отдельно был предложен профиль 4:2:2, который в иерархию не входит.

5.9.3. *MPEG-1* и *MPEG-2*

Документация *MPEG-1* была опубликована первой. В ней описываются стандарты кодирования с малыми цифровыми потоками – до 1,5 Мбит/с. Чересстрочная развертка не поддерживалась, разрешающая способность была ограничена максимум до 352 x 288 отсчетов. Для стереосигналов можно было пользоваться потоками до 192 Кбит/с. Величина цифрового потока была выбрана, в основном, исходя из значения этого параметра в стандарте компакт-диска, поскольку преследовалась цель использовать *MPEG-1* в аудио-визуальных системах, таких, как КД-И и КД-ПЗУ.

Ясно, что применение стандарта *MPEG-1* в ТВ вещании крайне ограничено, и чтобы решить эту проблему был разработан стандарт *MPEG-2*. Этот стандарт поддерживает чересстрочную развертку и позволяет кодировать сигналы ТСЧ с потоком от 3 до 15 Мбит/с или сигналы ТВЧ с потоком от 15 до 30 Мбит/с. Возможности кодирования звуковых сигналов расширены, в частности теперь поддерживается система кругового звука. Поскольку стандарты *MPEG* имеют иерархическую структуру, декодер *MPEG-2* может декодировать сигналы *MPEG-1*, но не наоборот.

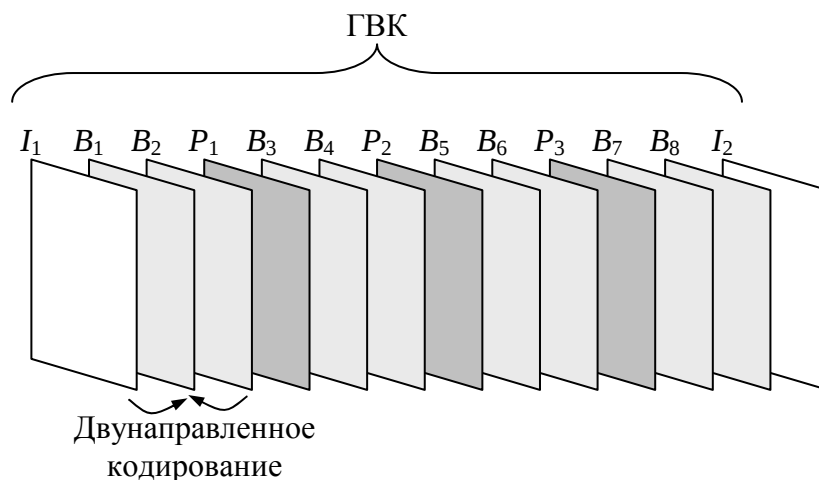
5.9.4. Двухнаправленное кодирование

Одной из самых сильных сторон *MPEG* является использование принципа двухнаправленного кодирования. Раздел 5 содержит введение в концепцию предсказания с компенсацией движения и *P*-видеокадров, которые декодируются с использованием опоры в виде предыдущего видеокadra, получившего необходимый сдвиг под воздействием векторов движения и скорректированного разностными данными. Двухнаправленное кодирование продвигает эту идею на шаг вперед, предусматривая возможность предсказания части текущего видеокadra как по предыдущим, так и по последующим видеокдрам.

Рис. 5.7 содержит иллюстративный пример последовательности видеокadres, используемой в стандарте *MPEG*. Последовательность начинается с

1-видеокадра, служащего опорой; совокупность этого и всех последующих видеокадров вплоть до появления очередного *I*-видеокадра называется «группой видеокадров» (ГВК). В составе ГВК имеется несколько видеокадров, полученных с прямым («вперед») предсказанием, или *P*-видеокадров. Первый такой видеокадр декодируется с использованием *I*-видеокадра в качестве опорного. Следующий и остальные *P*-видеокадры декодируются с использованием опоры в виде предыдущего *P*-видеокадра. Остальные видеокадры в ГВК – это *B*-видеокадры. Последние могут декодироваться с использованием данных, расположенных рядом, до или после; возможно также использовать данные, усредненные по этим двум видеокадрам. Поскольку все видеокадры сгруппированы в макроблоки, к каждому макроблоку может быть применено либо прямое либо обратное («назад») предсказание. Очевидно, что обратное предсказание невозможно, если декодер еще не получил видеокадра, который необходим для этого обратного предсказания. Решение состоит в передаче видеокадров в неправильном порядке. Вслед за *I*-видеокадром посылается первый *P*-видеокадр. Как только декодер получает *I*- и *P*-видеокадры, появляется возможность декодировать промежуточные *B*-видеокадры, для чего данные, полученные после компенсации движения, следует переместить вперед или назад. Очевидно, что система двунаправленного кодирования требует иметь больший объем памяти в кодере и декодере, а также вносит при кодировании увеличенную задержку. Ради уменьшения сложности в простом профиле *MPEG* использование *B*-видеокадров не предусматривается, поэтому при заданном цифровом потоке он не способен обеспечить такое же хорошее качество.

Полезность двунаправленного кодирования проясняется при рассмотрении границ движущегося объекта. Со стороны переднего края изображение фона не просматривается, поэтому для будущих кадров потребуется меньший объем данных. Данные фонового плана уже известны декодеру, поэтому прямое



Порядок передачи: $I_1 \ P_1 \ B_1 \ B_2 \ P_2 \ B_3 \ B_4 \ P_3 \ B_5 \ B_6 \ I_2 \ B_7 \ B_8$

Рис. 5.7. – Последовательность видеокадров (стандарт *MPEG*)

предсказание наиболее эффективно. Однако со стороны заднего края объекта фон открывается лишь при движении, и данные фонового изображения в предшествующих кадрах не содержатся. Однако в последующих кадрах фон открывается все больше и больше, поэтому эффективнее воспользоваться данными более позднего кадра, частично их отбрасывая. На практике кодер сжатия может производить кодирование как с прямым, так и с обратным предсказанием, однако посылаться будут только разностные данные одного из них, имеющие наименьший объем.

5.9.5. Типы данных

В результате цифрового сжатия одного видеосигнала согласно стандартам *MPEG* должны быть получены данные нескольких типов, перечисленных ниже:

1. Декодеру следует сообщить профиль, уровень и размер видеокадра, выраженный числом макроблоков, хотя эти сведения не требуются часто повторять.
2. Должна быть описана последовательность *I*-, *P*- и *B*-видеокадров, чтобы декодеру было известно, каким образом интерпретировать и изменить порядок следования данных передаваемых видеокадров.
3. Время, к которому относится видеокадр, указывается посредством передачи разновидности временного кода.
4. Поскольку в процессе цифрового сжатия используется переменное взвешивание, может оказаться необходимым передать значения весовой матрицы, если только не используется значение по умолчанию.
5. Собственно видеоданные, позволяющие декодировать видеокадры трех указанных типов.

5.9.6. Структура цифрового потока *MPEG*

Самым верхним компонентом структуры одиночного цифрового потока *MPEG*, как показано на рис. 5.8, является «видеопоследовательность». Эта последовательность включает в себя несколько ГВК. Заголовок последовательности специфицирует профиль и уровень, размер видеокадра и частоту кадров. Дополнительно можно передавать значения используемой весовой матрицы. Несколько ГВК, входящих в состав видеопоследовательности, передаются по очереди. Каждая ГВК начинается с заголовка ГВК, содержащего временную метку первого видеокадра (в виде временного кода).

В каждой ГВК первым всегда передается *I*-видеокадр, а затем могут следовать *P*- или *I*-видеокадры. Заголовок видеокадра содержит кодовое слово, специфицирующее тип видеокадра и его положение в ГВК. Поскольку в случае двунаправленного кодирования видеокадры посылаются в неправильном

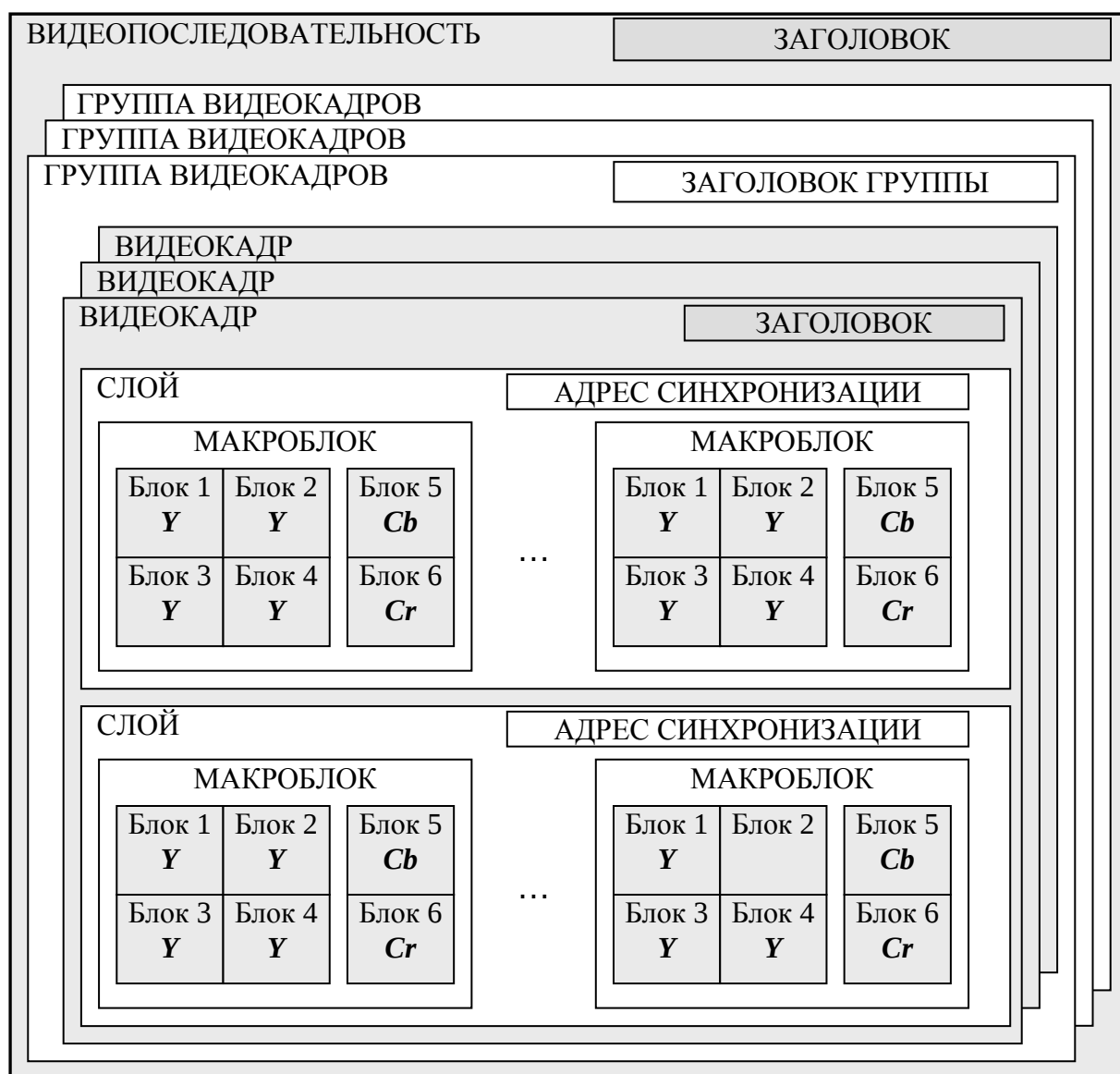


Рис. 5.8. – Синтаксис цифрового потока *MPEG2* (без расширений)

порядке, заголовок видеокadra играет важную роль в восстановлении правильного порядка для отображения.

Каждый видеокادر разбивается по двум координатам на блоки размером 8 x 8 отсчетов. Поскольку цветоразностный сигнал ранее был подвергнут субдискретизации с коэффициентом 2:1, на каждые четыре блока яркостных отсчетов приходится только один блок отсчетов сигнала *R-Y* и один блок отсчетов сигнала *B-Y*. Набор из четырех яркостных блоков, образующих квадрат, и двух соответствующих цветоразностных блоков представляет участок видеокadra, данные которого переносятся макроблоком. Макроблок – это единица информации о видеокadre, для которой предусматривается один вектор движения; таким образом, в составе каждого макроблока будет передаваться двумерный вектор, за исключением *I*-видеокadres, которые, по определению, не имеют векторов движения. В *P*-видеокadres векторы соответ-

вуют прямому предсказанию, а в случае *B*-видеокадров требуется указать, какому предсказанию соответствует вектор – прямому, обратному или комбинированному на основе этих двух видов.

Ряд макроблоков по ширине экрана называется слоем. Условие равенства протяженности слоя ширине экрана не имеет существенного значения, однако вертикальная граница видеокadra должна совпадать с границей слоя. В случае сигнала яркости по Рекомендации 601 и слоя, занимающего всю ширину экрана, одна строка содержит 720 отсчетов, то есть слой состоит из 45 макроблоков, поскольку $45 \times 16 = 720$. Каждый слой в составе видеокadra содержит индивидуальный адрес для целей синхронизации. Совокупность нескольких слоев составляет полный видеокادر. Например, видеокادر согласно низкому профилю содержит максимум 288 строк. Видеокادر будет содержать 18 слоев, так как $18 \times 16 = 288$.

5.9.7. Системный уровень

Структура, которая описывалась до сих пор, известна как элементарный поток, поскольку она является носителем данных только одного видеоканала. Другие элементарные потоки переносят данные звукового сопровождения. Следующий уровень кодирования касается мультиплексирования видео- и аудиоданных нескольких программ. Речь идет о так называемом системном уровне.

Программный поток составляется из видео- и аудиоданных одной вещательной службы, имеющих одну общую тактовую частоту. Программные потоки имеют пакеты данных переменной длины, которые относительно плохо защищены от ошибок, и поэтому они пригодны только для систем передачи или записи, не имеющих ошибок.

Транспортный поток является альтернативой программного потока. В нем используются пакеты постоянной длины с коррекцией ошибок. На рис. 5.9 показан пакет транспортного потока, включающий 187 байтов данных. Хотя стандарты *MPEG* не требуют этого в явной форме, общепринято защищать такой пакет от ошибок посредством добавления избыточных 16 байтов. Системный уровень предоставляет также таблицы распределения программ, в которых перечисляются все передаваемые элементарные потоки и вещательные службы, к которым эти потоки относятся. Поскольку службы могут пользоваться различными источниками тактовых импульсов, предусматривается их независимое тактирование.

Синхрослово	Данные, 187 байт	Проверочные байты, 16 байт
-------------	------------------	----------------------------

Рис. 5.9. – Пакет транспортного потока

Применение кодирования с переменной длиной слова означает, что происходит динамическое увеличение или уменьшение содержимого буферного ЗУ данных в декодере и кодере, результатом чего становится изменение задержки при кодировании и декодировании. В системном уровне предусмотрены временные метки, интегрированные в данные, которые гарантируют синхронность декодированных сигналов звука и изображения, поступающих на выход. Если применяется условный доступ, системный уровень обеспечивает управление скремблированием и шифрованием.

ЛИТЕРАТУРА

1. Гольденберг Л.М., Матюшкин Б.Д., Поляк М.Н. Цифровая обработка сигналов: Учебное пособие для вузов. – М.: Радио и связь, 1990. – 256 с.
2. Карташев В.Г. Основы теории дискретных сигналов и цифровых фильтров: Учебное пособие для вузов. – М.: Высшая школа, 1982. – 109 с.
3. Казанцев Г.Д., Курячий М.И., Пустынский И.Н. Измерительное телевидение: Учебное пособие для вузов. – М.: Высшая школа, 1994. – 288 с.
4. Оппенгейм А.В., Шафер Р.В. Цифровая обработка сигналов: Пер. с англ. – М.: Связь, 1979. – 416 с.
5. Рабинер Л., Гоулд Б. Теория и применение цифровой обработки сигналов: Пер. с англ. – М.: Мир, 1978. – 848 с.
6. Даджион Д., Мерсеро Р. Цифровая обработка многомерных сигналов: Пер. с англ. – М.: Мир, 1988. – 488 с.
7. Антонию А. Цифровые фильтры: анализ и проектирование: Пер. с англ. – М.: Радио и связь, 1983. – 320 с.
8. Каппелини В., Константинодис А., Эмилиани П. Цифровые фильтры и их применение. – М.: Энергоатомиздат, 1983. – 360 с.
9. Оппенгейм А. Применение цифровой обработки сигналов: Пер. с англ. – М.: Мир, 1980. – 552 с.
10. Остапенко А.Г., Сушков А.Б., Бутенко В.В. Рекурсивные цифровые фильтры на микропроцессорах. – М.: Радио и связь, 1988. – 128 с.
11. Куприянов М.С., Матюшкин Б.Д. Цифровая обработка сигналов: процессоры, алгоритмы, средства проектирования. – 2-е изд. – СПб.: Политехника, 1999. – 592 с.
12. Гоноровский И.С., Демин М.Н. Радиотехнические цепи и сигналы: Учебное пособие для вузов. – 5-е изд., перераб. и доп. – М.: Радио и связь, 1994. – 480 с.
13. Уоткинсон Дж. Пособие для инженеров по цифровому сжатию. – М.: «Снелл и Уилкокс», 1994. – 64 с.