

ФЕДЕРАЛЬНОЕ АГЕНТСТВО СВЯЗИ

**Федеральное государственное
образовательное бюджетное учреждение
высшего профессионального образования
«САНКТ-ПЕТЕРБУРГСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ТЕЛЕКОММУНИКАЦИЙ
им. проф. М. А. БОНЧ-БРУЕВИЧА»**

А.В.Никитина, А.Е.Рыжков

**ЦИФРОВАЯ ОБРАБОТКА СИГНАЛОВ В СЕТЯХ
РАДИОДОСТУПА**

УЧЕБНОЕ ПОСОБИЕ

СПбГУТ)))

**САНКТ-ПЕТЕРБУРГ
2013**

УДК 621.396.93(075.8)
ББК 38.84.16я73
Б12

Рецензенты

*Утверждено редакционно-издательским советом СПбГУТ
в качестве учебного пособия*

Никитина А.В.,

Б12 Цифровая обработка сигналов в сетях радиодоступа/
А.В.Никитина, А.Е.Рыжков – СПб.: Редакционно-издательский
центр СПбГУТ, 2013. – с.

Рассмотрены процедуры обработки цифровых сигналов в сетях радиодоступа на физическом уровне. Представлены структуры и алгоритмы работы вокодеров, используемых в сетях стандартов GSM/UMTS. Обсуждены вопросы реализации методов избыточного кодирования в сетях радиодоступа, построения адаптивных эквалайзеров в приемной аппаратуре стандарта GSM.

Предназначено для студентов, обучающихся по направлению подготовки 210700 квалификации “бакалавр” и квалификации “магистр”, а также специалистов в области телекоммуникаций.

**УДК 621.396.93(075.8)
ББК 38.84.16я73**

© Никитина А.В., Рыжков А.Е., 2013

© Федеральное государственное образовательное бюджетное учреждение высшего профессионального образования «Санкт-Петербургский государственный университет телекоммуникаций им. проф. М. А. Бонч-Бруевича», 2012

**Никитина Александра Викторовна
Рыжков Александр Евгеньевич**

Цифровая обработка сигналов в сетях радиодоступа

Учебное пособие

Редактор Л.К. Паршина
Верстка Н.А. Ефремовой

План 2013 г., п. 70

Подписано к печати
Объем 5,25 усл.-печ. л. Тираж 130 экз. Заказ 425

Редакционно-издательский центр СПбГУТ
191186 СПб., наб. р. Мойки, 61
Отпечатано в СПбГУТ

ВВЕДЕНИЕ

В технике радиосвязи 21 века большинство процедур, связанных с обработкой сигналов, формированием и приемом радиосигналов, выполняют программно процессорные блоки. Программно осуществляют сжатие и синтез речевых сигналов, избыточное кодирование, перемежение, скремблирование, шифрацию блоков передаваемых данных и обратные преобразования при приеме. Базовые алгоритмы обработки сигналов в сетях мобильной связи на системном уровне рассмотрены в [1]. Их конкретной реализации была посвящена монография [2], вышедшая в 1999г. За прошедшие 15 лет процессорная техника далеко ушла вперед. Это позволило реализовать высокоскоростные сети радиодоступа с кодовым разделением каналов и сети с технологией OFDM на физическом уровне. Произошли значительные изменения и в алгоритмах обработки сигналов, которые должны быть представлены в учебной литературе.

Пособие состоит из 3 частей. В главе 1 рассмотрены современные вокодеры, используемые в сетях GSM/UMTS при передаче речи с коммутацией каналов. В качестве базового взят основной используемый вокодер AMR (Adaptive Multi-Rate) класса ACELP. В русскоязычной литературе он практически не описан, в том числе и в учебном пособии [3]. Представлен также алгоритм работы детектора активности речи VAD (Voice Activity Detector) для вокодера AMR.

Глава 2 посвящена вопросам применения избыточного кодирования в сетях радиодоступа. Теория и алгоритмы кодирования рассмотрены в учебной литературе весьма подробно. Из последних изданий стоит отметить учебное пособие [4]. Поэтому в главе 2 основное внимание сконцентрировано на конкретных технологиях кодирования, используемых в сетях радиодоступа.

В главе 3 рассмотрены адаптивные эквалайзеры, работающие в приемниках абонентских и базовых станций сетей GSM. В учебной литературе на русском языке эти эквалайзеры не описаны.

Материалы учебного пособия соответствуют программе курса для магистров направления 210700 “Цифровая обработка сигналов в сетях радиодоступа” и разделам курса “Физические основы систем мобильной связи” для бакалавров направления 210700 профессионально-образовательных программ “Системы мобильной связи”.

1. ТЕХНОЛОГИИ ПЕРЕДАЧИ ТЕЛЕФОНИИ В СЕТЯХ ПОДВИЖНОЙ СВЯЗИ

1.1. Основы цифровой передачи телефонии

Цифровую передачу речи используют для повышения качества передачи, поскольку цифровая передача обеспечивает высокую помехоустойчивость каналов связи. Блок-схема кодера показана на рис. 1.1:

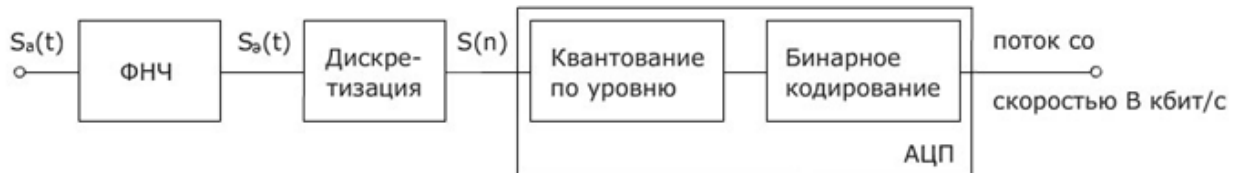


Рис.1.1. Блок-схема кодера

Фильтр нижних частот ограничивает полосу частот сигнала, чтобы избежать наложения спектров после дискретизации.

Полоса $P_{\text{ФНЧ}} = 300 \dots 3400$ Гц

Процедура дискретизации во времени показана на рис.1.2.

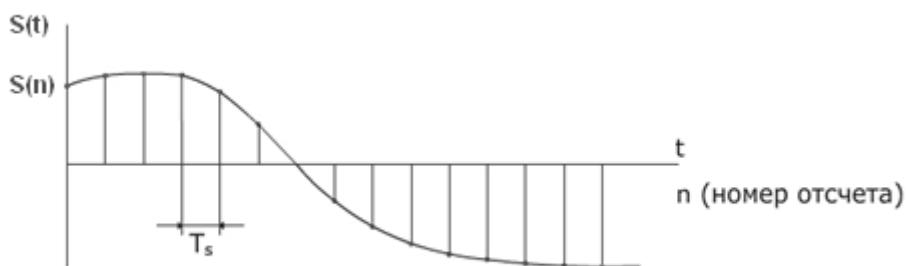


Рис.1.2. Дискретизация сигнала во времени

T_s – период дискретизации, для телефонии $T_s = 125$ мкс.

$f_s = 1/T_s$ – частота дискретизации 8 кГц ($f_s \geq 2F_{\text{max}}$)

Спектр $G(f)$ сигнала $S(n)$ (решетчатой функции) после дискретизации по времени показан на рис.1.3.

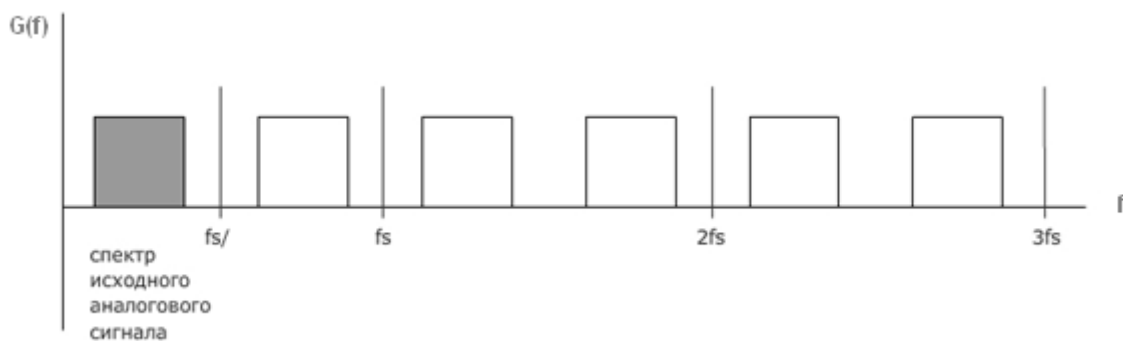


Рис.1.3. Спектр сигнала после дискретизации

Процедуру квантования: замену отсчетов решетчатой функции в вольтах на целые числа (рис.1.4) выполняет АЦП.

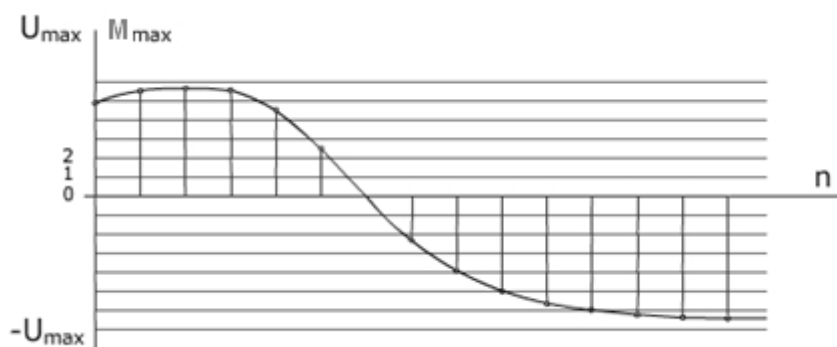


Рис.1.4. Квантование дискретного сигнала

При бинарном кодировании отсчеты представлены в виде двоичных чисел, а каждому отсчету соответствует m -разрядное двоичное число (рис.1.5). В старшем разряде записан знак: 0 при “+” и 1 при “-”.



Рис.1.5. Формат передаваемых чисел

При передаче стандартного телефонного потока $M_{max}=127$, $m=8$ и скорость передачи

$$B = m \times f_s = 8 \times 8 \times 10^3 = 64 \text{ кбит/с}$$

Требуемая полоса основных частот для передачи

$$P_{осн} \geq \frac{B}{2} = 32 \text{ кГц} \quad (1.1)$$

В радиоканале при использовании двухпозиционной модуляции необходима полоса

$$P_{радио} \geq B = 64 \text{ кГц} \quad (1.2)$$

При квантовании возникают помехи из-за дискретности квантователя. Интервал квантования (расстояние между двумя соседними уровнями) $\Delta = \frac{2U_{max}}{2^m}$, где m – разрядность отсчетов. Ошибка x распределена в этом интервале равномерно. Плотность вероятности распределения x :

$$W(x) = \frac{1}{\Delta}$$

Дисперсия x

$$D(x) = \int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} x^2 W(x) dx = \int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} x^2 \frac{1}{\Delta} dx = \frac{\Delta^2}{12} = \frac{U_{max}^2}{3 \cdot 2^{2m}}$$

$D(x)$ – это мощность шумов квантования. Найдем отношение сигнал/шум SNR (signal to noise ratio).

$$SNR = \frac{P_{сигн}}{D(x)} = \frac{u^2(t)}{U_{max}^2} \cdot \frac{U_{max}^2}{D(x)} = 3 \cdot 2^{2N} \left[\frac{u(t)}{U_{max}} \right]^2$$

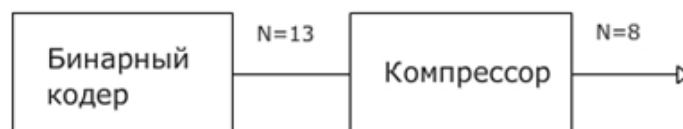
Результат, представленный в децибелах:

$$SNR(\text{дБ}) = 10 \lg \left[\frac{P_{сигн}}{D(x)} \right] = 4,8 + 6N - 20 \log \left(\frac{U_{max}}{u(t)} \right)$$

При максимальном уровне сигнала $u(t) = U_{max}$ и $m = 8$, отношение сигнал/шум $SNR = 4,8 + 48 = 52,8$ дБ. Примерно такой же уровень (не менее 50 дБ) требуется и при передаче слабых сигналов: 3 – 5% от пикового. Но при $m = 8$ и малых значениях $u(t)$ SNR будет хуже на $20 \lg(0,03) = -30$ дБ.

Увеличение m на один разряд приводит к увеличению отношения сигнал/шум на 6 дБ. Чтобы получить увеличение SNR на 30 дБ необходимо увеличить разрядность на $30/6=5$ разрядов, то есть вместо 8 разрядов кодировать 13. Однако при этом возрастет скорость передачи речи и требуемая полоса канала связи. Чтобы обеспечить передачу 8-разрядными числами в телефонной сети общего пользования (ТФОП) прибегают к сжатию динамического диапазона речевого сигнала при передаче (компрессии) и его восстановлению при приеме (экспандировании). Эти операции выполняет программное устройство – компрессор (рис.1.6).

При передаче



При приеме

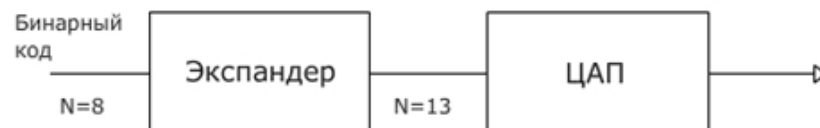


Рис.1.6. Структура компрессора

Компрессия основана на использовании логарифмического закона при сжатии сигнала. На практике применяют два варианта компрессии:

- американский: на основе μ -закона,
- европейский: на основе A -закона.

Для μ -закона

$$y = \text{sign} [\bar{s}] U_{max} \frac{\ln(1 + \mu|\bar{s}|)}{\ln(1 + \mu)}$$

Где $\bar{s} = \frac{u(n)}{U_{max}}$; $\mu = 255$.

Для А-закона:

$$y = \begin{cases} \frac{A \cdot s}{1 + \ln A}, |\bar{s}| \leq \frac{1}{A} \\ \text{sign}(s) U_{max} \frac{1 + \ln(A|\bar{s}|)}{1 + \ln A}, \frac{1}{A} \leq |\bar{s}| \leq 1 \end{cases}$$

$A=87,56$ при $m=8$, $A=86$ при $m=7$.

Характеристики компрессоров для А- и μ -законов показаны на рис.1.7.

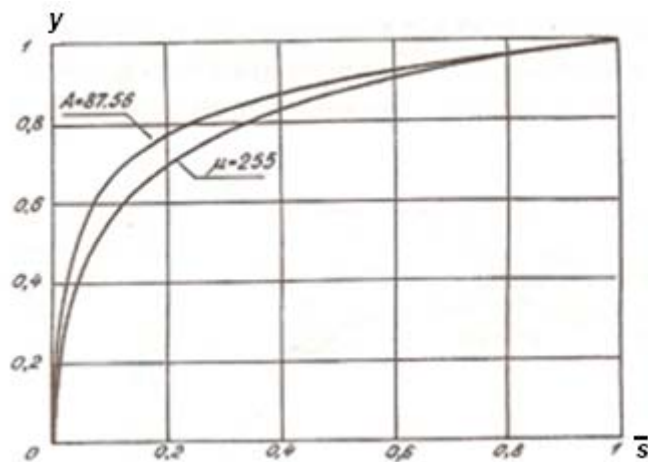


Рис.1.7. Законы компрессии речи

1.2. Снижение скорости передачи речи. Дифференциальная ИКМ

Как следует из (1.1) и (1.2), за высокое качество передачи сигналов в цифровом виде приходится платить многократным увеличением полосы канала связи в сравнении с полосой исходного информационного сигнала. Поэтому, особенно в радиоканалах, актуальна задача уменьшения требуемой полосы, для чего используют методы сжатия цифрового информационного сигнала. Наиболее простым способом является использование дифференциального кодирования, когда вместо самого сигнала $s(n)$ передают разность (ошибку) между $s(n)$ и его предсказанным значением $\tilde{s}(n)$:

$$e(n) = s(n) - \tilde{s}(n) \quad (1.3)$$

Схема кодека дифференциальной ИКМ (ДИКМ) показана на рис.1.8.

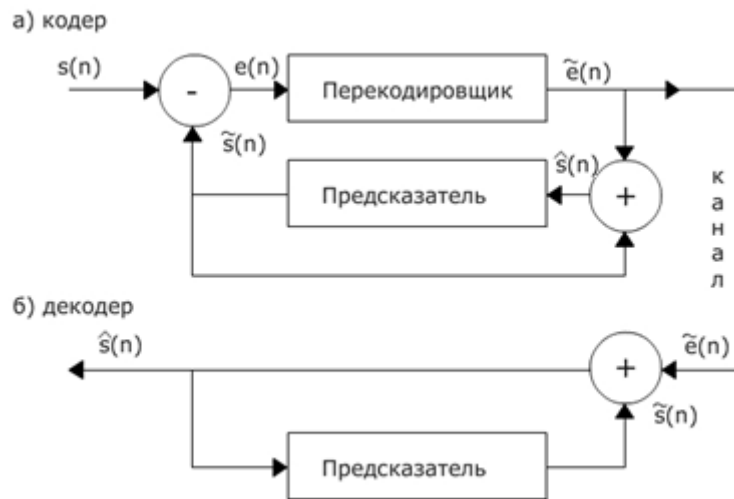


Рис. 1.8. Блок-схема ДИКМ

В работе предсказателей используют алгоритм линейного предсказания, когда предсказываемое значение сигнала представляют в виде линейной суммы взвешенных предыдущих отсчетов:

$$\tilde{s}(n) = a_1 s(n-1) + a_2 s(n-2) + \dots + a_p s(n-p) \quad (1.4)$$

Обычно $p \leq 8 \dots 10$, так как при дальнейшем увеличении p точность предсказания не повышается.

Тогда:

$$e(n) = s(n) - \sum_{k=1}^p a_k s(n-k) \quad (1.5)$$

Восстановленное значение отсчета

$$\hat{s}(n) = \tilde{s}(n) + \tilde{e}(n).$$

Обычно ошибки $e(n)$ оказываются на полпорядка ниже самих отсчетов $s(n)$, что позволяет заметно уменьшить число разрядов отсчетов $e(n)$ в сравнении с разрядностью $s(n)$. При $m=8$ скорость передачи речевого потока падает в 2 раза (32 кбит/с, вместо 64 кбит/с).

Основной проблемой при ДИКМ является оптимизация коэффициентов предсказателя $\{a_k\}$, так как от их выбора зависят ошибки $e(n)$. Оптимизацию выполняют на основе минимизации функции среднеквадратичной ошибки:

$$E = \sum_{n=n_1}^{n=n_2} e^2(n) = \sum_{n=n_1}^{n=n_2} [s(n) - \sum_{k=1}^p a_k s(n-k)]^2, \quad (1.6)$$

где число отсчетов p между процедурами адаптации соответствует интервалу в 5...7мс между отсчетами n_1 и n_2 . На практике используют разные варианты решения задачи минимизируют $E[a_1, a_2, \dots, a_p]$. Приведем простейший вариант, когда адаптация коэффициентов $\{a_k\}$ происходит на каждом шаге. Это градиентный метод.

Запишем уравнение для квадрата ошибки $e^2(n)$ в следующем виде:

$$e^2(n) = [s(n) - \bar{a}(k) \bar{s}^T(n-k)]^2 \quad (1.7)$$

где вектор $\bar{a} = \{a_1, a_2, \dots, a_p\}$, а вектор $\bar{s}(n-k) = \{s(n-1), s(n-2), \dots, s(n-p)\}$

Изменение коэффициентов $\{a_k\}$ производят по следующему закону:

$$\bar{a}^{(m+1)} = \bar{a}^{(m)} - [\nabla e^2(n)] \bar{a}^{T(m)} \quad (1.8)$$

В (1.6) $\bar{a}^{(m+1)}$ – улучшенное значение \bar{a} , ∇ – градиент функции $e^2(n)$ по вектору \bar{a} :

$$[\nabla e^2(n)] = 2e(n)[-s^T(n-k)] \quad (1.9)$$

Тогда:

$$\bar{a}^{(m+1)} = \bar{a}^{(m)} + 2e(n)\bar{a}^{(m)}s^T(n-k) \quad (1.10)$$

Дифференциальную ИКМ с адаптивно меняющимися коэффициентами $a(k)$ называют адаптивной ДИКМ (АДИКМ). АДИКМ находит применение в простых системах радиотелефонии, например, в DECT, где $B = 32$ кбит/с в радиоканале.

1.3. Дельта-модуляция

При Δ -модуляции используют ту же схему, что при ДИКМ (рис.1.8), но разрядность передачи ошибки $e(n)$ составляет 1 бит. В результате по каналу связи передают последовательность вида: показанного на рис.1.9.



Рис.1.9. Сигнал при Δ -модуляции

Положительной ошибке соответствует передача +1 (логического 0), а отрицательной ошибки -1 (логической 1). Поскольку структура передачи не изменилась, то Δ -модуляция позволяет снизить скорость передачи в сравнении с ИКМ в 2 раза, как и ДИКМ. Однако, в отличие от ДИКМ для передачи того же качества речи требуется увеличивать частоту выборок с 8кГц до 32кГц, то есть в 4 раза.

Проблемой при использовании Δ -модуляции являются ошибки, возникающие при передаче постоянных уровней (рис.1.10)

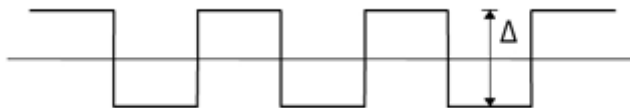


Рис.1.10. Передача постоянного уровня при Δ -модуляции

и быстрых переходов (рис. 1.11).

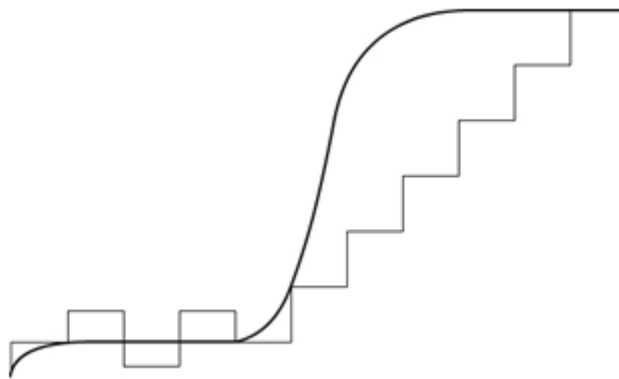


Рис.1.11. Ошибки при передаче быстрых переходов

В первом случае (рис.1.10) следует уменьшить Δ , во втором (рис.11) – увеличить. Для этого используют адаптивную Δ -модуляцию, где величину Δ меняют по определенному закону. Например, при передаче постоянных уровней Δ мала. Однако, если подряд следуют исключительно логические нули или логические единицы, то на каждом следующем шаге величину Δ увеличивают в 2 раза до того момента, пока знак производной функции не изменится. Тогда опять $\Delta = \Delta_{\min}$ и процесс изменения Δ начинается заново.

1.4 Эффективное кодирование речи. Принципы сжатия речевых сигналов.

Работа над стандартами сотовой связи 2-го поколения потребовала создания новых технологий цифровой передачи речи по радиоканалам. В результате были разработаны кодеры-декодеры речевого сигнала, называемые **вокодерами** (сокращение от английских слов voice coder), позволившие снизить скорость информационного речевого сигнала в радиоканале ниже 13 кбит/с без заметного ухудшения качества речи в телефоне. Работа вокодеров основана на устранении избыточности, содержащейся в речевом сигнале, путём анализа сигнала при передаче и его воссоздания при приеме. В результате по каналу связи передают не собственно речь в цифровом виде, а ее характеристики, позволяющие **синтезировать речевой сигнал** в приемнике. Прежде чем перейти к рассмотрению вокодеров, следует дать краткую характеристику самих речевых сигналов, а также органов человека, участвующих в формировании и восприятии речи. Эти органы универсальны для всех людей, а анализ речевых сигналов показывает идентичность их основных свойств для различных языков.

Схема голосового аппарата показана на рис.1.12. В формировании речи участвуют легкие, трахеи, гортань, рот и нос. Гортань содержит две складки кожи, называемые голосовыми связками, между которыми расположена голосовая щель. Голосовые связки регулируют поток воздуха, проходящий через них, периодически открывая и закрывая голосовую щель. Таким образом, воздух в рот поступает в виде последовательности импульсов, частоту следования которых называют частотой **основного тона** (pitch frequency).

Пройдя голосовую щель, поток воздуха попадает в рот, представляющий собой нелинейную акустическую систему, длина которой составляет приблизительно 17 см у взрослого мужчины. При изменении положения губ, языка, челюстей и перепонки меняется объем и конфигурация ротовой полости, а, следовательно, и резонансные частоты акустической системы. Перепонка регулирует

связь между ртом и носом. При произнесении неносовых звуков перепонка перекрывает вход в нос, и воздух проходит только через рот. При произнесении носовых или смешанных звуков перепонка открывает вход воздуха в нос, который также является нелинейной акустической системой. Длина носовой полости мужчины приблизительно равна 12 см. Во время речи легкие при дыхании наполняются воздухом. Выдыхаемый по трахее через голосовую щель поток воздуха и является первичным источником речи. Голосовой аппарат, воздействуя на проходящий сквозь него воздух, формирует различные звуки, которые принято различать на звонкие, глухие и взрывные.



Рис.1.12. Схема голосового аппарата человека

При формировании звонких звуков (гласных и согласных) голосовые связки колеблются, пропуская поток воздуха в виде квазипериодической последовательности. Эти последовательности содержат много гармоник частоты основного тона, затухающих приблизительно по закону 12 дБ/октава. Голосовой аппарат действует как резонансная система, усиливая одни из этих гармоник и подавляя другие. Говорящий управляет частотой основного тона только в известных пределах. Так, у взрослого мужчины частота основного тона меняется в пределах от 50 до 250 Гц, у взрослой женщины от 120 до 300 Гц и до 350 Гц у детей.

Как было сказано, голосовой тракт является сложной перестраиваемой резонансной системой. В простейшем случае, при произнесении нейтральных (безударных) гласных, его можно представить в виде трубы длиной L , с одной стороны которой в узле колебаний расположен источник звука (голосовые связки), а другая, где находится пучность звуковых волн, открыта (раскрытые губы). Такая труба резонирует на нечетных гармониках собственной частоты: $f_0, 3f_0, 5f_0$. Частота $f_0 = v_{зв}/4L$, где $v_{зв} = 340$ м/с – скорость звука. При длине голосового тракта 17 см получаем $f_0 = 500$ Гц и другие резонансные частоты: 1500 Гц, 2500 Гц и т.д. Эти резонансные частоты в акустике называют **формантами**. В процессе речи частоты формант постоянно меняются.

Спектр **звонких звуков** в основном расположен в нижней полосе частот речевого сигнала и сгруппирован вокруг формант, однако скорость изменений

формант такова, что спектр речи остается практически постоянным на промежутках менее 16 мс.

Глухие звуки представляют собой слабые шумоподобные колебания. При формировании глухих звуков голосовые связки не колеблются, а в голосовом тракте возникает сжатие, проходя которое поток воздуха генерирует широкий спектр частот (шум). В зависимости от того, где возникает сжатие (в передней части рта или в районе голосовых связок), воздействие голосового тракта на формируемый шум различно. Однако во всех случаях в спектре глухих звуков подчеркнуты верхние частоты.

При произнесении **взрывных звуков** голосовой тракт перекрыт в каком-то месте и затем, после мгновенного сжатия препятствия в нем возникает переходный акустический процесс. Если к этому добавляется вибрация голосовых связок, то формируются звонкие взрывные звуки (например, твёрдые согласные д, м и т.п.).

Слуховая система человека (ухо) устроена так, что воспринимает весь спектр звуковых частот набором из большого количества биологических полосовых фильтров, настроенных по всему диапазону звуковых частот. При восприятии речи проявляются эффекты маскирования отдельных тонов и сигналов. Так, сильные звуки подавляют слабые, что позволяет отбрасывать их при передаче без существенного ущерба для качества звучания восстановленного сигнала. Кроме того, отдельные частоты могут маскироваться сигналами близкоотстоящих от них частот. Динамический диапазон речевых сигналов для различных людей достигает 60 дБ при изменении уровня речи отдельного человека в пределах 20 - 30 дБ.

Рассмотренные кратко механизмы образования и восприятия речи позволяют сформулировать основные идеи, использованные в технике анализа-синтеза речи при ее передаче в сжатом виде по каналам связи. Речь можно синтезировать с помощью генератора периодических импульсов, следующих с частотой основного тона для воспроизведения звонких звуков, и генератора шума для воспроизведения глухих звуков. Оба исходных сигнала следует пропустить через фильтры с меняющимися во времени характеристиками, соответствующими характеристикам речевого тракта.

Речевой тракт содержит большое число резонансных частот или формант. Для воспроизведения речи на уровне узкополосной телефонной связи достаточно ограничиться первыми тремя-четырьмя формантами, расположенными в диапазоне 100 – 3500 Гц. Амплитуды более высоких формант затухают по закону 12дБ/октава, но для высококачественного воспроизведения речи (широкополосной телефонии) требуется расширять полосу воспроизводимых частот от 70 Гц до 7 кГц.

Устранение избыточности в информационных сигналах основано на принципе: **можно не передавать то, что можно предсказать**. А предсказать можно периодические процессы. Так, вместо передачи синусоиды по точкам, можно передать по каналу ее амплитуду и частоту, и при приеме эту синусоиду синтезировать.

В речи выделяют те части, которые на определенных временных участках близки к периодическим процессам, и передают их не в виде множества отсчетов, а описательно, с помощью набора параметров. При передаче речевой сигнал подвергают анализу, удаляя из него избыточность, обусловленную формантами и возбуждением звонких звуков импульсами, следующими с частотой основного то-

на. Для удаления формант и основного тона используют кодирование с линейным предсказанием. В результате по каналу связи передают шумовой сигнал возбуждения и ряд независимых параметров, позволяющих синтезировать в приемнике речевой сигнал, максимально близкий к исходному.

Важнейшей характеристикой вокодера является качество восстанавливаемого (синтезируемого) сигнала. Наряду с разработанными инструментальными методами оценки качества речи широко используют субъективный метод на основе **усредненных экспертных оценок** по пятибалльной (иногда десятибалльной) шкале, где оценка 5 (отлично) определяет наивысшее качество передачи. Если для речи, ограниченной полосой телефонного канала 300 - 3400 Гц, качество соответствует оценке 5, то для междугородной телефонной связи оно должно соответствовать оценке 4. Для систем подвижной связи и радиодоступа оценка речевого сигнала должна лежать в диапазоне 3,5 – 4. При её снижении до 2,5 – 3,0 речь, сохраняя хорошую разборчивость, теряет естественность звучания, становится "синтетической", неиндивидуализированной. Такая передача возможна в специальных системах связи (служебные, военные и т.п.), но недопустима в коммерческих системах телефонной связи. Поэтому вокодеры, применяемые в сотовых телефонах, сложны как по структуре, так и по заложенному в них программному обеспечению. Следует также учесть, что усложнение обработки речевого сигнала в вокодерах может производиться только за счет увеличения скорости обработки, так как при телефонии в реальном времени допустимые задержки сигнала по всему тракту передачи не должны превышать 150 мс.

1.5. Обработка речи в вокодерах

Процесс кодирования речи в вокодере можно разбить на ряд типовых процедур, каждую из которых выполняют методами цифровой фильтрации. При кодировании выявляют периодические процессы в речевом сигнале, определяют их параметры, а затем устраняют из речевого сигнала избыточность, исключая найденные периодичности. В результате получают остаточный речевой сигнал, который аппроксимируют и передают вместе с параметрами периодических процессов в речи по каналу связи. Перечисленные процедуры составляют процесс анализа речи. В приемнике, в декодирующем устройстве, по принятому остаточному сигналу и параметрам периодических процессов речи восстанавливают речевой сигнал, выполняя процесс синтеза речи.

Основные процедуры анализа речевого сигнала $u(t)$ показаны на рис.1.13.

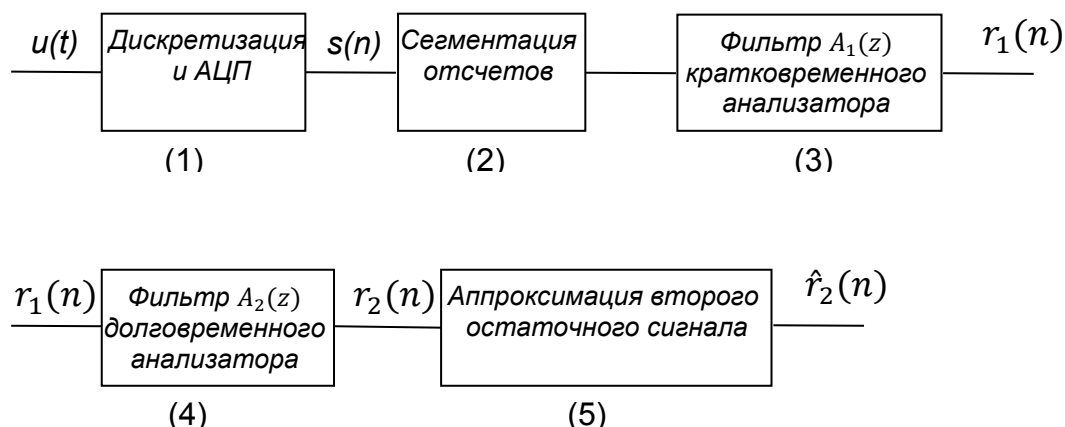


Рис.1.13. Процедуры анализа речевого сигнала

1. После дискретизации речевого сигнала (при обычной узкополосной телефонии с частотой выборок $f_s = 8$ кГц и соответственно периодом выборок $T_s = 125$ мкс) выборки аналогового сигнала с помощью АЦП преобразуют в цифровой код, над которым и совершают все последующие операции. Каждой выборке соответствует двоичное число длиной в 13 бит.

2. Последовательность отсчетов разбивают на отдельные, следующие друг за другом сегменты. Длительность сегмента выбирают так, чтобы на его протяжении состояние речевого тракта мало менялось. Естественно, чем короче сегмент, тем с большей точностью его можно описать набором характерных параметров, но при этом общее количество параметров, передаваемых в единицу времени, возрастает и, следовательно, увеличивается скорость передачи. В большинстве кодеков речи систем подвижной связи речевой сигнал разбивают на сегменты длительностью 20 мс. При этом каждый сегмент содержит $20 \times 10^{-3} \times 8 \times 10^3 = 160$ отсчетов речевого сигнала.

3. В каждом отдельном сегменте выявляют периодичность, обусловленную формантами речи. Эту операцию выполняют с помощью процедуры, называемой кратковременным линейным предсказанием, а устройство, ее реализующее, называют кратковременным линейным предсказателем.

Идея метода состоит в том, что каждый отсчет $s(n)$ внутри отдельного сегмента аппроксимируют линейной комбинацией из p предшествующих отсчетов с неизменными для данного сегмента коэффициентами a_k :

$$\tilde{s}(n) = \sum_{k=1}^p a_k s(n - k). \quad (1.11)$$

В (1.11) $\tilde{s}(n)$ – оценка или приближенное значение отсчета, которое отличается от исходного $s(n)$ на величину ошибки

$$e(n) = s(n) - \tilde{s}(n) = s(n) - \sum_{k=1}^p a_k s(n - k). \quad (1.12)$$

Выполняя z-преобразование над обеими частями этого выражения, получаем

$$e(z) = A_1(z) \cdot s(z), \quad (1.13)$$

где

$$A_1(z) = 1 - \sum_{k=1}^p a_k z^{-k}. \quad (1.14)$$

В предсказателе для каждого сегмента определяют коэффициенты цифрового фильтра с передаточной характеристикой $A_1(z)$ (Здесь и далее будем использовать при описании цифровой обработки сигналов во времени функции вида $F(n)$, где n – номер текущего отсчёта, и обозначение $F(z)$ при переходе к z-преобразованию).

Если $s(n)$ – цифровой речевой сигнал, а $s(z)$ – его z-преобразование, то устранение из речевого сигнала избыточности, обусловленной колебаниями на формантах, производят с помощью фильтрации:

$$r_1(z) = A_1(z)s(z). \quad (1.15)$$

называют первым остаточным сигналом. Зная его, можно восстановить (синтезировать) исходный сигнал, выполняя обратное преобразование

$$s(z) = r_1(z)A_1^{-1}(z) = r_1(z)H_1(z) \quad (1.16)$$

где $H_1(z) = 1/A_1(z)$ – передаточная функция цифрового фильтра-синтезатора, инверсная $A_1(z)$ фильтра-анализатора. Функция $H_1(z)$ аппроксимирует огибающую сегмента речи и содержит одни полюсы. Число p определяет порядок фильтра и его называют порядком предсказателя, а коэффициенты $a_k (k \in 1 \dots p)$ – коэффициентами предсказателя. Эти коэффициенты одинаковы для обоих фильтров.

Вычисление вектора коэффициентов \bar{a}_k является целью первого этапа обработки сегмента. Оба цифровых фильтра: и анализатора, и синтезатора, являются адаптивными, так как их параметры (коэффициенты) меняются от сегмента к сегменту.

Отметим, что выражения (1.12) и (1.15) совпадают с формулой (1.5) при ДТКМ. Следовательно, снижение скорости передачи речи при ДИКМ происходит за счет удаления формант.

4. На втором этапе анализа выявляют тонкую структуру речи, а именно периодичность на частоте основного тона. Для этого используют процедуру долговременного линейного предсказания, в результате чего появляется возможность устранить избыточность (периодичность) из первого остаточного сигнала с помощью цифровой фильтрации

$$r_2(z) = A_2(z)r_1(z), \quad (1.17)$$

где $A_2(z)$ – передаточная функция фильтра-анализатора долговременного предсказателя, а $r_2(n)$ – ошибка долговременного предсказания, или второй остаточный сигнал. Восстановить первый остаточный сигнал можно, используя фильтр-синтезатор долговременного предсказателя с передаточной функцией

$$H_2(z) = 1/A_2(z);$$

$$r_1(z) = H_2(z)r_2(z). \quad (1.18)$$

5. Третий этап анализа - аппроксимация второго остаточного сигнала с целью максимального уменьшения числа бит, необходимых для его передачи. Второй остаточный сигнал $r_2(n)$ – некоррелированный шум. В вокодерах, применявшихся в конце 20 века, его аппроксимировали, например, путем прореживания отсчетов. В вокодерах 21 века с кодовым возбуждением вместо $r_2(n)$ передают квазишумовой сигнал возбуждения $\tilde{r}_2(n)$, который выбирают таким образом, чтобы при синтезе речи в приемнике получился сигнал $\tilde{s}(n)$, минимально отличающийся от исходного. Введем функцию $K(z)$, связывающую реальный второй остаточный сигнал (1.17) и сигнал возбуждения $\tilde{r}_2(z)$:

$$\tilde{r}_2(z) = K(z)r_2(z) \quad (1.19)$$

При синтезе речи в приемнике в соответствии с (1.16), (1.18) и (1.19) получим

$$\tilde{s}(z) = \frac{K(z)r_2(z)}{A_1(z)A_2(z)} = K(z)s(z) \quad (1.20)$$

Следовательно, исходный сигнал $s(z)$ отличается от синтезированного сигнала $\hat{s}(z)$ на величину ошибки $e_s(z) = s(z)(1-K(z))$. Естественное желание минимизировать ошибку $e_s(n)$ приводит к структуре речевого кодера с петлей управления (обратной связью) (рис.1.14). Такую структуру называют **кодером по схеме анализ-синтез**.

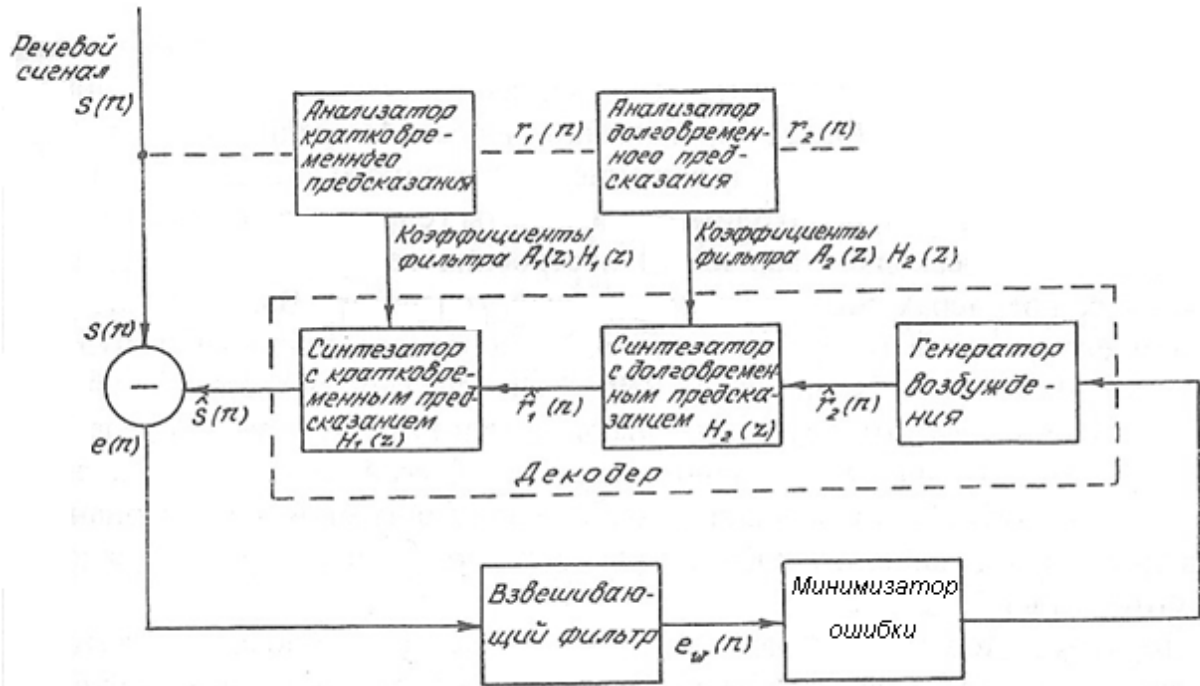


Рис.1.14. Структура кодирующего устройства по схеме анализ-синтез

В этой схеме на этапах анализа определяют коэффициенты фильтров кратковременного $H_1(z)$ и долгосрочного $H_2(z)$ предсказателей, а затем производят поиск оптимального сигнала возбуждения $\hat{r}_2(n)$ минимизирующего ошибку $e(n) = s(n) - \hat{s}(n)$ между исходным и синтезированным сигналами. Ошибку оценивает управляющий блок (минимизатор ошибки) на основе среднеквадратичного критерия минимизации. Перед этим блоком стоит взвешивающий фильтр. Поскольку на частотах формант шумы маскированы более сильными звучаниями формант, то ошибки на этих частотах менее заметны для слушателей. Взвешивающий фильтр снижает уровень сигнала ошибки на частотах формант, поднимая уровни других частот сигнала ошибки.

В состав кодера поностью входит декодер приемника, синтезирующий речевой сигнал в соответствии с (1.20). В отличие от декодера в кодере операция (1.20) повторяется множество раз при поиске оптимального сигнала возбуждения $\hat{r}_2(n)$.

1.6. Вокодер AMR

Основой вокодер, используемый в сетях GSM/UMTS при передаче телефони с коммутацией каналов – это многоскоростной адаптивный вокодер AMR (Adaptive Multi-Rate). Вокодер позволяет производить сжатие речевых сигналов до скоростей 12,2; 10,2; 7,95; 7,4; 6,7; 5,9; 5,15 и 4,74 кбит/с [5]. Вокодер специфицирован в 2001г. и является развитием вокодера EFR (Enhanced Full Rate), разработанного в 1995г.

Вокодер AMR относится к классу вокодеров с кодовым возбуждением ACELP (Algebraic Code Excited Linear Prediction). При удалении формант и основного тона используют процедуры линейного предсказания; код возбуждения формируют с помощью алгебраического алгоритма поиска.

Кодер AMR построен по структуре анализ-синтез (рис.1.14) с небольшими изменениями. Так как фильтрация является линейной операцией, то вместо взвешивающего фильтра в цепи сигнала ошибки $e(n)$ введены 2 фильтра в обеих ветвях сигналов $s(n)$ и $\hat{s}(n)$. При этом они совмещены с синтезирующими фильтрами формант $H_1(z)$. В результате получаем следующую функциональную схему кодера AMR (рис. 1.15):

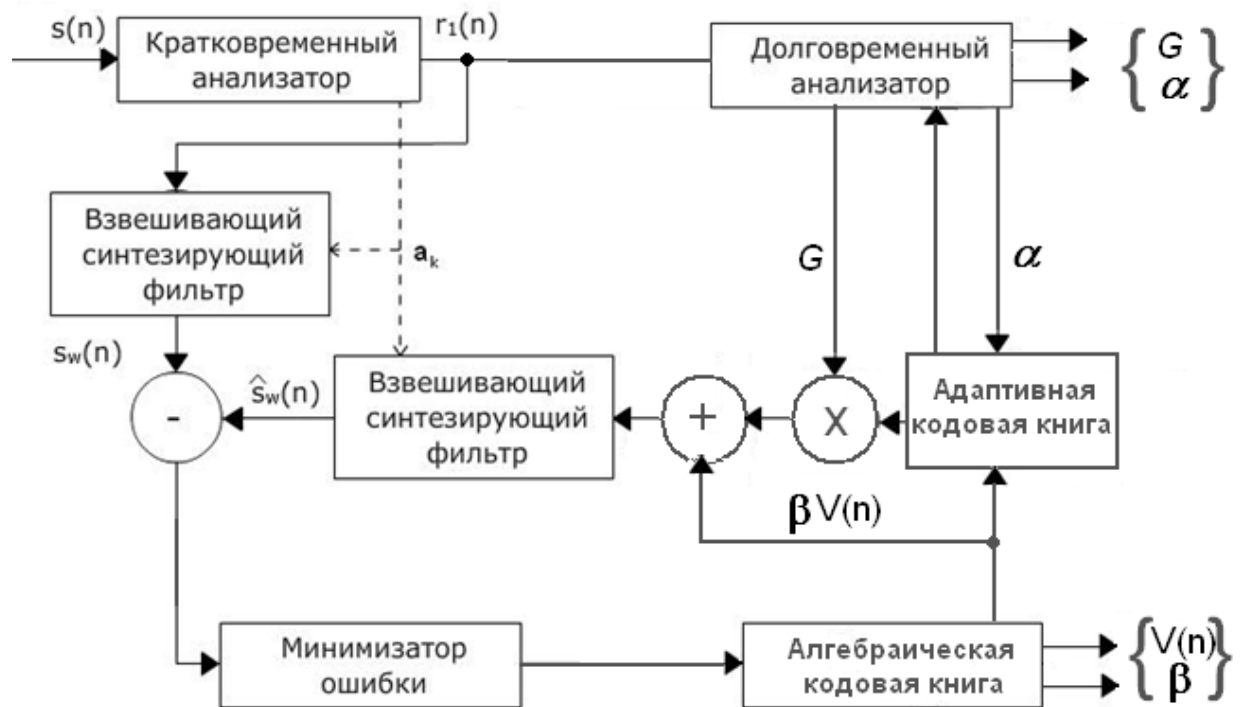


Рис.1.15. Функциональная схема кодера AMR

Характеристики взвешивающего фильтра выбирают на основе характеристик синтезирующего фильтра кратковременного предсказателя с дополнительными коэффициентами γ_1 и γ_2 :

$$W(z) = \frac{1 - \sum_{k=1}^p a_k \gamma_1^k z^{-k}}{1 - \sum_{k=1}^p a_k \gamma_2^k z^{-k}}, \quad (1.21)$$

где $0 \leq \gamma_2 < \gamma_1 \leq 1$. При скоростях $B=12,2$ и $10,2$ кбит/с $\gamma_1 = 0,9$ и $\gamma_2 = 0,6$; при меньших скоростях $\gamma_1 = 0,94$ и $\gamma_2 = 0,6$. Передаточная функция фильтра (1.21) определяет его характеристики. Если $\gamma_1 = \gamma_2$, то коэффициент передачи $W(z) = 1$ и взвешивания сигнала ошибки нет. Если же $\gamma_2 = 0$, то взвешивающий фильтр работает как фильтр-анализатор $A_1(z)$ в кратковременном анализаторе.

В схеме 1.15 характеристики фильтра кратковременного анализатора

$$A(z) = 1 - \sum_{k=1}^p a_k z^{-k}, \quad (1.22)$$

а взвешивающего синтезирующего фильтра

$$H_w(z) = H(z) \cdot W(z) = \frac{W(z)}{A(z)} = \frac{\left(1 - \sum_{k=1}^p a_k \gamma_1^k z^{-k}\right)}{\left(1 - \sum_{k=1}^p a_k z^{-k}\right) \cdot \left(1 - \sum_{k=1}^p a_k \gamma_2^k z^{-k}\right)}, \quad (1.23)$$

Взвешенный входной сигнал

$$s_w(n) = s(n) - \sum_{k=1}^p a_k \gamma_1^k s(n-k) + \sum_{k=1}^p a_k \gamma_2^k s(n-k), \quad (1.24)$$

а минимизируемая взвешенная среднеквадратичная ошибка.

$$E_w = \sum_{n=0}^{N-1} e_w^2(n) = \sum_{n=0}^{N-1} [s_w(n) - \hat{s}_w(n)]^2 \quad (1.25)$$

Для сжатия речи в вокодере AMR цифровой речевой сигнал, состоящий из 13-битовых отсчетов речи, следующих с частотой выборок 8 кГц, разбивают на сегменты длиной 20 мс. В свою очередь каждый сегмент делят на 4 подсегмента по 5 мс каждый. В сегменте – 160 выборок (отсчетов), в подсегменте – 40.

Процесс сжатия сегмента речи состоит из 3-х процедур.

1. Определение коэффициентов фильтров удаления $A_1(z)$ и синтеза $H_1(z)$ формант и удаление формант.
2. Определение коэффициентов фильтра $H_2(z)$ синтеза основного тона.
3. Формирование сигнала возбуждения и оптимизация его параметров.
4. Формирование пакета данных для передачи по радиоканалу: параметров фильтров $H_1(z)$, $H_2(z)$ и параметров сигнала возбуждения.

Рассмотрим последовательно эти процедуры.

Процедура определения коэффициентов \bar{a}_k фильтров $A_1(z)$ и $H_1(z)$ основана на минимизации среднеквадратичной ошибки между исходным $s(n)$ и предсказанным сигналами $\tilde{s}(n)$:

$$E = \sum_n e_1^2(n) = \sum_n \left[s(n) - \sum_{k=1}^p \bar{a}_k s(n-k) \right]^2, \quad (1.26)$$

В вокодере AMR $p=10$, так что система (1.30) состоит из 10 уравнений. Левая матрица системы (1.30) является симметричной матрицей Теплица и для вычисления вектора \bar{a}_k используют рекурсивный алгоритм Левинсона-Дурбина, специфицированный в [5].

В качестве начального условия выбирают начальную ошибку $E_{LD} = R(0)$, после чего следует итерационный процесс из 10 циклов (по числу коэффициентов a_k).

```

for i = 1 to 10 do
   $a_0^{(i-1)} = 1$ 
   $k_i = - \left[ \sum_{j=0}^{i-1} a_j^{(i-1)} R(i-j) \right] / E_{LD}(i-1)$ 
   $a_i^{(i)} = k_i$ 
  for j = 1 to i - 1 do
     $a_j^{(i)} = a_j^{(i-1)} + k_i a_{i-j}^{(i-1)}$ 
  end
   $E_{LD}(i) = (1 - k_i^2) E_{LD}(i-1)$ 
end

```

При этом на каждом i -м шаге появляется очередной коэффициент a_i и происходит уточнение коэффициентов с меньшими i .

Для эффективной передачи значений коэффициентов \bar{a}_k по радиоканалу на их основе рассчитывают линейные спектральные пары характеристик фильтров $A_1(z)$ и $H_1(z)$. Далее определяют взвешенные разности найденных векторов спектральных частот с их предсказанными значениями из результатов расчетов предыдущего сегмента речи, которые кодируют минимально возможным количеством бит. Объем передаваемой информации о спектральных парах в зависимости от скорости передачи сжатой речи приведен в конце параграфа в табл. 1.3.

Для того, чтобы при синтезе речи в приемнике учесть погрешность, вызванную кодированием информации о линейных спектральных парах, производят обратную процедуру расчета линейных спектральных пар по их закодированным значениям и определяют коэффициенты \bar{a}_k с учетом внесенных погрешностей. Эти округленные значения \bar{a}_k и используют в фильтрах $A_1(z)$ и $H_1(z)$.

В реальном алгоритме работы вокодера все выглядит сложнее.

Прежде всего надо обеспечить плавность переходов синтезируемого сигнала от одного 20мс сегмента к другому. Поэтому при расчете автокорреляционных функций (1.29) к 160 отсчетам текущего сегмента добавляют 80 последних отсчетов предыдущего сегмента. Итак, общее число отсчетов при вычислениях $R(j)$ составляет 240.

Второй важный момент состоит в том, что синтез речевого сигнала в приемнике выполняют не для всего 20мс сегмента, а последовательно для четырех 5мс подсегментов. При этом для каждого подсегмента используют свой вектор коэффициентов \bar{a}_k . Делают это следующим образом.

Автокорреляционные функции $R(j)$ рассчитывают для взвешенных значений $s'(n) = W(n) \times s(n)$. Графики весовых коэффициентов $W_I(n)$ и $W_{II}(n)$ показаны на рис. 1.16. В результате автокорреляционные функции (1.29) вычисляют по формуле

$$R(j) = \sum_{n=j}^{239} s'(n) \times s'(n-j), \quad j = 0 \dots 10. \quad (1.29')$$

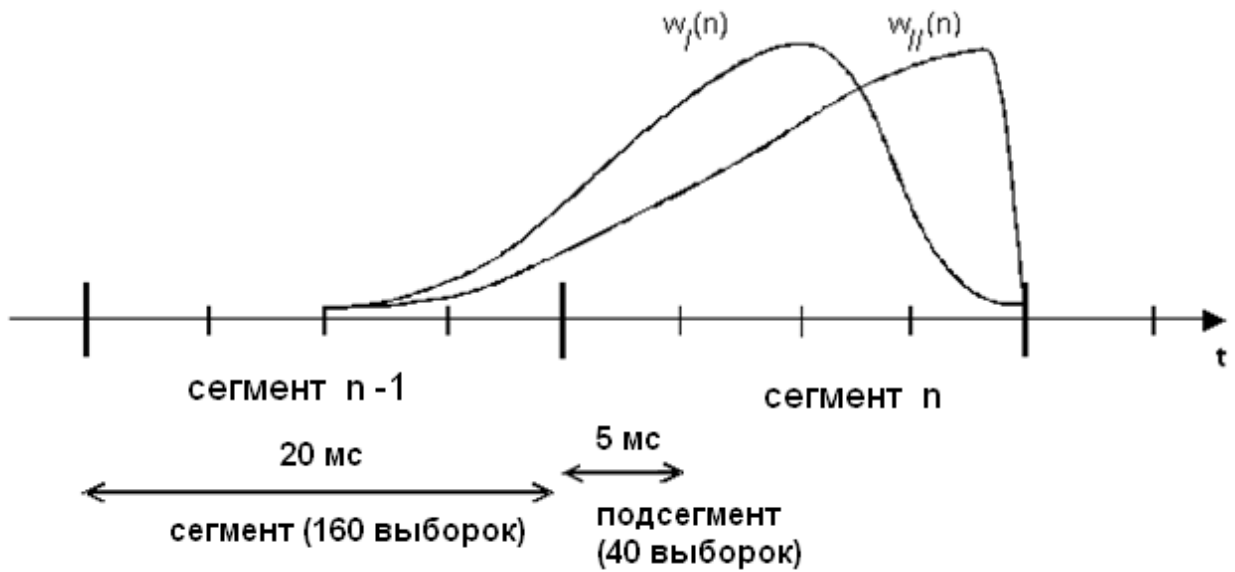


Рис.1.16. Весовые коэффициенты $W_I(n)$ и $W_{II}(n)$

Далее функцию $R(0)$ умножают на коэффициент 1,0001, что эквивалентно добавке шумового фона на уровне -40 дБ. Остальные функции $R(j)$ ($j = 1 \dots 10$) умножают на коэффициент

$$w_{lag}(j) = \exp \left[\frac{1}{2} \left(\frac{2\pi f_0 j}{f_s} \right)^2 \right],$$

где $f_0 = 60$ Гц, а $f_s = 8000$ Гц – частота выборки. Соответственно при вычислении вектора \bar{a}_k в систему (1.30) подставляют взвешенные функции: $R'(0) = 1,0001R(0)$ и $R'(j) = w_{lag} \times R(j)$.

При скорости передачи речи $B = 12,2$ кбит/с векторы коэффициентов \bar{a}_k рассчитывают дважды. При $s'(n) = W_I(n) \times s(n)$ – это будут коэффициенты \bar{a}_k для подсегмента №2, а при $s'(n) = W_{II}(n) \times s(n)$ – для подсегмента №4. Для подсегментов №1 и №3 векторы \bar{a}_k определяют после интерполяции значений линейных спектральных пар, рассчитанных для подсегментов №2 и №4 данного сегмента и №4 предыдущего сегмента. Для всех остальных скоростей передачи речи, от $B = 10,2$ кбит/с и ниже, расчет вектора \bar{a}_k выполняют один раз, только для $s'(n) = W_{II}(n) \times s(n)$. Полученный вектор используют для расчетов в подсегменте №4. Для остальных подсегментов (№1, №2, №3) векторы \bar{a}_k определяют путем интерполяции значений линейных спектральных пар, полученных для данного и предыдущего 20мс сегментов речи.

Два следующих этапа работы кодера: определение коэффициентов фильтра $H_2(z)$ синтеза основного тона и формирование сигнала возбуждения, касаются тонких структур речи, сильно влияющих на качество синтезируемого сигнала в

приемнике. Ниже рассмотрены базовые алгоритмы соответствующих процедур вокодера AMR, которые, кроме того, содержат много дополнительных деталей, необходимых для обеспечения высоких качественных характеристик передачи речи [5].

На этапе определения параметров основного тона выявляют колебания с частотой основного тона. Для этого, как и при определении формант, используют линейное предсказание, но поскольку основной тон по частоте значительно ниже формант, эту процедуру называют долговременным предсказанием (Long Term Prediction).

В долговременном предсказателе анализируют первый остаточный сигнал, то-есть, сигнал, из которого удалены форманты:

$$r_1(n) = s(n) - \sum_{k=1}^p a_k s(n-k) \quad (1.31)$$

Долговременный предсказатель характеризуют 2 параметра: задержка α и коэффициент передачи G . Эти параметры определяют для каждого 5мс подсегмента речи. В основе работы долговременного предсказателя лежит следующая идея выявления периодичности. Сдвигают 40 отсчетов сигнала $r_1(n)$ подсегмента по временной оси влево до тех пор, пока не будет найдена последовательность из 40 предшествующих отсчетов, максимально похожая на данную (рис. 1. 17).

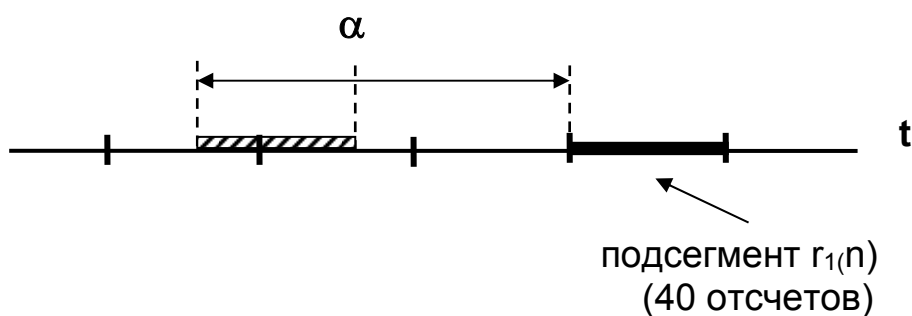


Рис.1.17. К определению сдвига α в долговременном предсказателе

Ошибка на выходе долговременного предсказателя представляет собой второй остаточный сигнал

$$e_2(n) = r_2(n) = r_1(n) - Gr_1(n-\alpha), \quad (1.32)$$

Выполняя над обеими частями (1.32) z-преобразование, получаем

$$r_2(z) = (1 - Gz^{-\alpha})r_1(z) \quad (1.33)$$

где $1 - Gz^{-\alpha}$ – передаточная характеристика цифрового фильтра с одним отводом. Фильтр с обратной передаточной характеристикой

$$H_2(z) = \frac{1}{1 - Gz^{-\alpha}} \quad (1.34)$$

позволяет восстановить в синтезаторе речи первый остаточный сигнал $r_1(n)$:

$$r_1(n) = r_2(n) + Gr_1(n-\alpha) \quad (1.35)$$

Выбор оптимальных G и α производят путем минимизации среднеквадратичной ошибки

$$E = \sum_{n=0}^{N_1-1} e_2^2(n) = \sum_{n=0}^{N_1-1} [r_1(n) - Gr_1(n-\alpha)]^2 \quad (1.36)$$

где N_1 – число анализируемых отсчетов. Приравнявая $\partial E/\partial G=0$, находим оптимальный коэффициент передачи

$$G = \frac{\sum_{n=0}^{N_1-1} r_1(n)r_1(n-\alpha)}{\sum_{n=0}^{N_1-1} [r_1(n-\alpha)]^2} \quad (1.37)$$

После подстановки (1.37) в (1.36) получим

$$E = \sum_{n=0}^{N_1-1} [r_1(n)]^2 - \frac{\left\{ \sum_{n=0}^{N_1-1} [r_1(n)r_1(n-\alpha)] \right\}^2}{\sum_{n=0}^{N_1-1} [r_1(n-\alpha)]^2} \quad (1.38)$$

Величину α находят перебором, максимизируя второе слагаемое в выражении (1.38), после чего вычисляют согласно (1.37) оптимальное G .

В вокоде AMR процесс определения параметров основного тона $\{G, \alpha\}$ состоит из двух процедур. Первая – начальный поиск α в “открытой петле”. Этот неточный поиск суживает зону поиска α , а также позволяет определить: идет передача речевого сигнала или следует пауза. Последнее необходимо для реализации технологии прерывистой передачи речи (см. п.1.8.).

Начальный поиск α выполняют, вычисляя нормированные автокорреляционные функции со сдвигом исходного речевого сигнала, прошедшего через взвешивающий фильтр:

$$R_k = \frac{\sum_{n=0}^{N'-1} s_w(n)s_w(n-k)}{\sum_{n=0}^{N'-1} s_w(n-k)s_w(n-k)} \quad (1.39)$$

В (1.39) $N'=80$. Максимуму R_k соответствует искомый сдвиг $\alpha=k$. При скоростях передачи речи $B = 12,2...5,9$ кбит/с эту процедуру выполняют дважды: для отсчетов $s_w(n)$ первого-второго и третьего-четвертого подсегментов. При $B = 5,15$ и $4,75$ кбит/с – один раз для каждого 20мс сегмента.

Во время второй процедуры определяют точные значения α для каждого подсегмента и соответствующие им коэффициенты передачи G . При этом в выражениях (1.37) и (1.38) вместо предшествующих отсчетов первого остаточного сигнала $r_1(n-\alpha)$ подставляют **отсчеты сигнала $r_{1синт}(n-\alpha)$, синтезированного в предыдущих подсегментах**, поскольку в приемнике существует только синтезированный сигнал. Соответственно (1.37) и (1.38) приобретают следующий вид:

$$G = \frac{\sum_{n=0}^{N_1-1} r_1(n)r_{\text{синт}}(n-\alpha)}{\sum_{n=0}^{N_1-1} [r_{\text{синт}}(n-\alpha)]^2} \quad (1.37')$$

$$E = \sum_{n=0}^{N_1-1} [r_1(n)]^2 - \frac{\left\{ \sum_{n=0}^{N_1-1} [r_1(n)r_{\text{синт}}(n-\alpha)] \right\}^2}{\sum_{n=0}^{N_1-1} [r_{\text{синт}}(n-\alpha)]^2} \quad (1.38')$$

Буфер синтезированного сигнала $r_{\text{синт}}(n-\alpha)$ в [] называют **адаптивной кодовой книгой**, а поиск α осуществляют в состоянии “замкнутой петли”. Величина N_1 в (1.38) составляет 40 (число отсчетов сигнала в подсегменте). Диапазон возможных смещений α лежит в пределах 19...143, что соответствует частотам основного тона 400...50 Гц.

Для повышения точности восстановления основного тона при задержках $\alpha < 95 \dots 85$ используют дробные числа, производя интерполяцию между ближайшими целочисленными значениями α . Так, при передаче с $B = 12,2$ кбит/с точность определения α в диапазоне задержек $[17+3/6; 96+3/6]$ составляет $1/6$ расстояния между отсчетами T_s . При $\alpha \geq 95$ точность ограничена целочисленной величиной T_s . Эту задержку определяют для подсегментов №1 и №3 и кодируют 9-разрядным двоичным числом (всего 512 возможных позиций). Для четных подсегментов (№2 и №4) определяют смещение $\Delta\alpha$ относительно задержки в нечетных подсегментах. Возможное смещение в диапазоне значений $\Delta\alpha$ $[-5+3/6; 4+3/6]$ кодируют 6-разрядным двоичным числом.

После определения α для каждого подсегмента рассчитывают G по формуле (1.37'). Коэффициент передачи G определяют в диапазоне $0 \leq G \leq 1,2$ и кодируют 4-разрядным двоичным числом.

При меньших скоростях передачи точность определения α составляет $(1/3)T_s$ в диапазоне значений $[19+1/3; 84+2/3]$; при $\alpha \geq 85$ передают целочисленные значения α . Число бит, выделяемых для передачи α в каждом подсегменте в зависимости от скорости передачи речи, приведено в табл. 1.3. Что касается коэффициента передачи G , то при $B = 7,95$ кбит/с его кодируют, как и при $B = 12,2$ кбит/с 4-х битовым числом. При других скоростях передачи коэффициент G кодируют совместно с коэффициентом передачи сигнала возбуждения.

Последний, 3-й этап – **синтез сигнала возбуждения** в вокодере типа ACELP состоит в поиске для каждого 5-мс подсегмента возбуждающего сигнала на основе **алгебраической кодовой книги**. В вокодерах с кодовым возбуждением второй остаточный сигнал $r_2(n)$ (1.32) отбрасывают, а вместо него генерируют такой сигнал возбуждения $\beta V(n)$, чтобы среднеквадратичная ошибка между взвешенным исходным сигналом и сгенерированным сигналом (1.25) была минимальной. При этом только небольшая часть отсчетов $V(n)$ отлична от нуля (в вокодере AMR при $B=12,2$ кбит/с 10 отсчетов из 40). Ненулевые отсчеты имеют одинаковую амплитуду, но знак их может быть любым. Таким образом, алгебраическая кодо-

вая книга является троичной: 1, -1, 0. Необходимо найти позиции ненулевых отсчетов $V(n)$ и коэффициент усиления β .

Синтезированный сигнал:

$$\hat{s}_w(n) = [\beta V(n) + Gr_{1cum}(n - \alpha)] ** h_w(n) + s_0(n) \quad (1.40)$$

В (1.40) знаком ** обозначена операция свертки во взвешивающем синтезирующем фильтре, а $s_0(n)$ – остаточный выходной сигнал от прохождения через фильтр предыдущего подсегмента. Ошибка

$$e(n) = s_w(n) - \hat{s}_w = s_w(n) - \{[\beta V(n) + Gr_{1cum}(n - \alpha)] ** h_w(n) + S_0(n)\} = x(n) - \beta V(n) ** h_w(n) \quad (1.41)$$

Величина $x(n) = s_w(n) - s_0(n) - Gr_{1cum}(n - \alpha) ** h_w(n)$ для каждого отсчета определена (фиксирована). Неизвестен сигнал $V(n)$, который надо выбрать. Целевая функция (среднеквадратичная ошибка):

$$E = \sum_{n=0}^{39} [x(n) - \beta V(n) ** h_w(n)]^2 \quad (1.42)$$

Найдем оптимальный коэффициент усиления β , приравнивая $\frac{\partial E}{\partial \beta} = 0$:

$$\beta = \frac{\sum_{n=0}^{39} x(n)[V(n) ** h_w(n)]}{\sum_{n=0}^{39} [V(n) ** h_w(n)]^2} \quad (1.43)$$

Числитель выражения для β представляет собой кросс-корреляционную функцию, а знаменатель – взвешенную энергию сигнала возбуждения.

Подставим (1.43) в (1.42):

$$E = \sum_{n=0}^{39} [x(n)]^2 - 2 \sum_{n=0}^{39} x(n) \frac{\sum_{n=0}^{39} x(n)[V(n) ** h_w(n)]}{\sum_{n=0}^{39} [V(n) ** h_w(n)]^2} [V(n) ** h_w(n)] + \frac{\left\{ \sum_{n=0}^{39} x(n)[V(n) ** h_w(n)] \right\}^2}{\left\{ \sum_{n=0}^{39} [V(n) ** h_w(n)]^2 \right\}^2} \sum_{n=0}^{39} [V(n) ** h_w(n)]^2 = \sum_{n=0}^{39} [x(n)]^2 - \frac{\left\{ \sum_{n=0}^{39} x(n)[V(n) ** h_w(n)] \right\}^2}{\sum_{n=0}^{39} [V(n) ** h_w(n)]^2} \quad (1.44)$$

Минимизация ошибки E требует максимизации второго слагаемого в (1.44). Обратим внимание на то, что при $V(n) ** h_w(n) = x(n)$ ошибка была бы равна нулю.

Вычисление второго слагаемого в (1.44) можно упростить. Для этого вспомним, как осуществляется операция свертки в цифровых фильтрах. На рис. 1.18а представлен некий входной сигнал $u(n)$, а на рис. 1.18б — импульсная характеристика фильтра $h(n)$.

Если $y(n) = u(n) ** h(n)$, то

$$y(0) = u(0) * h(0)$$

$$y(1) = u(0) * h(1) + u(1) * h(0)$$

$$y(2) = u(0) * h(2) + u(1) * h(1) + u(2) * h(0) \text{ и т.д.}$$

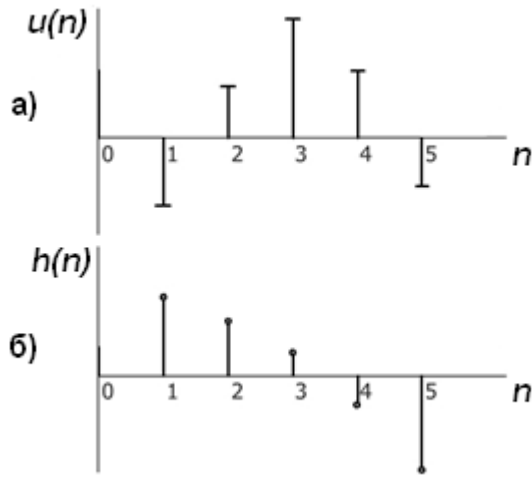


Рис.1.18. К определению операции свёртки

Представим операцию свертки в векторном виде.

$$\bar{u}(n) = \begin{bmatrix} u(0) \\ u(1) \\ \vdots \\ u(k) \end{bmatrix} \text{ – вектор-столбец, а свертка – умножение вектора-столбца на матрицу}$$

$$H_w = \begin{bmatrix} h(0) & & \\ h(1) & h(0) & \\ h(2) & h(1) & h(0) \end{bmatrix}$$

Матрица H_w – треугольная матрица Тевлица размерностью [40, 40], где по главной диагонали стоят $h(0)$, а в левом нижнем углу – $h_w(39)$. Тогда второе слагаемое в (1.44) можно записать так:

$$R(n) = \frac{[\bar{x}^T(n) * H_w * \bar{V}(n)]^2}{[\bar{V}^T(n) * H_w^T] * [H_w * \bar{V}(n)]} \quad (1.45)$$

Так как операции свертки и умножения коммутативны, то

$$R(n) = \frac{\{[\bar{x}^T(n) * H_w] * \bar{V}(n)\}^2}{[\bar{V}^T(n) * (H_w^T * H_w)] * \bar{V}(n)} \quad (1.46)$$

В числителе (1.46) выражение в квадратных скобках фиксировано, так что вычисление числителя сводится к перемножению двух векторов. Вычисление знаменателя сложнее, но матрицу $\Phi = (H_w^T * H_w)$ рассчитывают один раз, а вектор $V(n)$ содержит большое число нулей. В вокодере AMR используют алгоритмы, упрощающие вычисление $\bar{V}^T(n) * \Phi * \bar{V}(n)$.

Для поиска оптимальных 40 отсчетов $V(n)$ ($V(0) \dots V(39)$) их разбивают на несколько групп (треков). Число групп и число ненулевых отсчетов в подсегменте в зависимости от скорости передачи речи приведено в табл. 1.1.

Таблица 1.1

В (кбит/с)	Число групп (треков)	Число ненулевых отсчетов
12,2	5	10
10,2	4	8
7,95	4	4
7,4	4	4
6,7	3	3
5,9	2	2
5,15	2	2
4,75	2	2

Рассмотрим алгоритм поиска $V(n)$ при максимальной скорости передачи речи 12,2 кбит/с. 40 отсчетов вектора $V(n)$ разбиты на 5 групп, по 8 отсчетов в каждой (табл. 1.2).

Таблица 1.2

Группа	Номера отсчетов
1	0, 5, 10, 15, 20, 25, 30, 35
2	1, 6, 11, 16, 21, 26, 31, 36
3	2, 7, 12, 17, 22, 27, 32, 37
4	3, 8, 13, 18, 23, 28, 33, 38
5	4, 9, 14, 19, 24, 29, 34, 39

В результирующий вектор $V(n)$ попадут по два отсчета из каждой группы. Алгоритм поиска начинается с того, что для каждой группы, перебирая все 8, находят один отсчет, при котором $R(n)$ максимально. Из 5 найденных максимумов определяют глобальный максимум. Отсчет, соответствующий глобальному максимуму, обозначают как i_0 . Пусть в примере это будет отсчет 23 из группы 3.

На следующем этапе вычисляют $R(n)$ для 4 комбинаций $\{i_0, i_1\}$, где i_1 соответствует номеру отсчета локального максимума в группе. По максимуму $R(n)$ определяют i_1 (пусть ему соответствует отсчет 5 из группы 1) и таким образом, фиксируют пару $\{i_0, i_1\}$. Оставшиеся группы расставляют по убыванию $R(n)$ (пусть в примере: 5, 2, 4) и начинают искать последовательно пары $\{i_2, i_3\}$, $\{i_4, i_5\}$, $\{i_6, i_7\}$, $\{i_8, i_9\}$. В рассматриваемом примере i_2, i_3, i_4 – номера первых ненулевых отсчетов из групп 5, 2 и 4, а $i_5 \dots i_9$ – вторых ненулевых отсчетов соответствующих групп. Отсчеты каждой найденной пары фиксируют. При поиске пары $\{i_2, i_3\}$ возможны 8 позиций каждого отсчета в группе, что с учетом возможных 4 вариантов знаков от-

счетов требует 256 вычислений $R(n)$ для всех возможных комбинаций ненулевых отсчетов. При поисках следующих пар в тех группах, где уже фиксирован один отсчет, остается 7 возможных позиций для поиска. Таким образом на этом этапе просматривают $4 \times (8 \times 8 + 8 \times 7 + 7 \times 7 + 7 \times 7)$ комбинаций.

После того, как определен максимум $R(n)$ по всем итерациям, фиксируют значения отсчетов в каждой группе и знак. Если оба отсчета в группе имеют одинаковый знак, то их кодируют в порядке возрастания номера. Если знаки разные, то вначале кодируют отсчет с большим номером. Таким образом, для передачи отсчетов одной группы требуется 7 бит, а 5 групп – 35 бит. Коэффициент усиления β кодируют 5 битами. В результате для передачи параметров одного речевого 20мс сегмента требуется:

38 бит – коэффициенты фильтра кратковременного предсказания,
 $2 \times (9+4; 6+4) = 46$ бит – характеристики долговременного предсказателя,
 $4 \times (35+5) = 160$ бит – параметры сигнала возбуждения.

Всего: $160+38+46 = 244$ бита, что соответствует скорости передачи 12,2 кбит/с.

В табл.1.3 приведено число бит, используемых для кодирования отдельных параметров вокодера AMR и общее число передаваемых бит для 20мс сегмента речи в зависимости от ее скорости передачи в радиоканале

Таблица 1.3

Скорость передачи, кбит/с	Параметр	1-й под-сегмент	2-й под-сегмент	3-й под-сегмент	4-й под-сегмент	Всего
12,2 (GSM EFR)	Спектральн. пары					38
	Задержка α	9	6	9	6	30
	Кoeff. передачи G	4	4	4	4	16
	Алгебраич. книга	35	35	35	35	140
	Кoeff. передачи β	5	5	5	5	20
	Всего					244
10,2	Спектральн. пары					26
	Задержка α	8	5	8	5	26
	Алгебраич. книга	31	31	31	31	124
	Кoeff. передачи	7	7	7	7	28
	Всего					204
7,95	Спектральн. пары					27
	Задержка α	8	6	8	6	28

	Кэфф. пе- редачи G	4	4	4	4	16
	Алгебраич. книга	17	17	17	17	68
	Кэфф. пе- редачи β	5	5	5	5	20
	Всего					159
7,4	Спектральн. пары					26
	Задержка α	8	5	8	5	26
	Алгебраич. книга	17	17	17	17	68
	Кэфф. пе- редачи	7	7	7	7	28
	Всего					148
5,9	Спектральн. пары					26
	Задержка α	8	4	8	4	24
	Алгебраич. книга	11	11	11	11	44
	Кэфф. пе- редачи	6	6	6	6	24
	Всего					118
5,15	Спектральн. пары					23
	Задержка α	8	4	8	4	24
	Алгебраич. книга	9	9	9	9	36
	Кэфф. пе- редачи	6	6	6	6	24
	Всего					103
4,75	Спектральн. пары					23
	Задержка α	8	4	4	4	20
	Алгебраич. книга	9	9	9	9	36
	Кэфф. пе- редачи	8		8		16
	Всего					95

Процесс синтеза сигнала в декодере приемника иллюстрирует рис. 1.19.

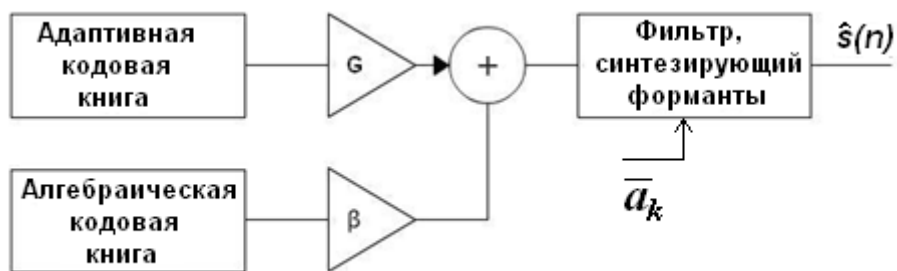


Рис. 1.19. Структура декодера

1.7. Развитие техники передачи речи в сотовых сетях связи

Для сетей стандарта GSM при его появлении в 1991г. были разработаны 2 типа вокодеров:

- полноскоростной (FR – Full Rate) со скоростью передачи телефонии в радиоканале 13 кбит/с [6],
- полускоростной (HR – Half Rate) со скоростью 5,6 кбит/с [7].

Полноскоростной вокодер построен по структуре RPE-LTP (Regular Pulse Excitation – Long Term Prediction): с возбуждением регулярной последовательностью импульсов с долговременным предсказанием. В 1991г. недостаточное быстродействие микропроцессоров не позволяло реализовывать вокодеры с кодовым возбуждением типа CELP и процедуры сжатия речи в вокодере RPE-LTP требует значительно меньшего числа операций.

Как и во всех вокодерах, в вокодере RPE-LTP процедура сжатия речевого сигнала состоит из 3-х этапов, но петель обратной связи в отличии от схемы (рис. 1.14) охвачены только 2 последние ступени: удаления основного тона из первого остаточного сигнала $r_1(n)$ и генерации сигнала возбуждения $r_{2синт}(n)$.

На первом этапе определяют коэффициенты фильтров удаления $A_1(z)$ и синтеза $H_1(z)$ формант в 20мс сегменте речи. Как и в вокодере ACELP, эту процедуру выполняют, минимизируя ошибку на основе линейного предсказания (1.26). В вокодере RPE-LTP число отсчетов кратковременного предсказателя $p=8$. По радиоканалу передают кодированные логарифмические отношения коэффициентов отражения k_i из алгоритма Левинсона-Дурбина, описанного в предыдущем параграфе 1.6. Для передачи данных о коэффициентах фильтра $H_1(z)$ требуется 36 бит.

После удаления формант получают первый остаточный сигнал $r_1(n)$. Блок-схема его обработки при передаче показана на рис. 1.20. Для удаления основного тона из первого остаточного сигнала необходимо найти параметры G и α в выражении (1.32). Эти параметры, как и в вокодере ACELP, находят для четырех 5-мс подсегментов каждого 20-мс сегмента речи. Для уменьшения ошибки, вызванной последующей аппроксимацией второго остаточного сигнала $r_2(n)$, в кодере вокодера RPE-LTP введена петля обратной связи, охватывающая 2-ю и 3-ю ступени анализа (рис.1.20). В результате, как и в вокодере ACELP, в поисковом буфере вместо отсчетов $r_1(n)$ предыдущих сегментов записаны отсчеты синтезированного первого остаточного сигнала и

$$r_2(n) = r_1(n) - G r_{1синт}(n - \alpha) \quad (1.47)$$

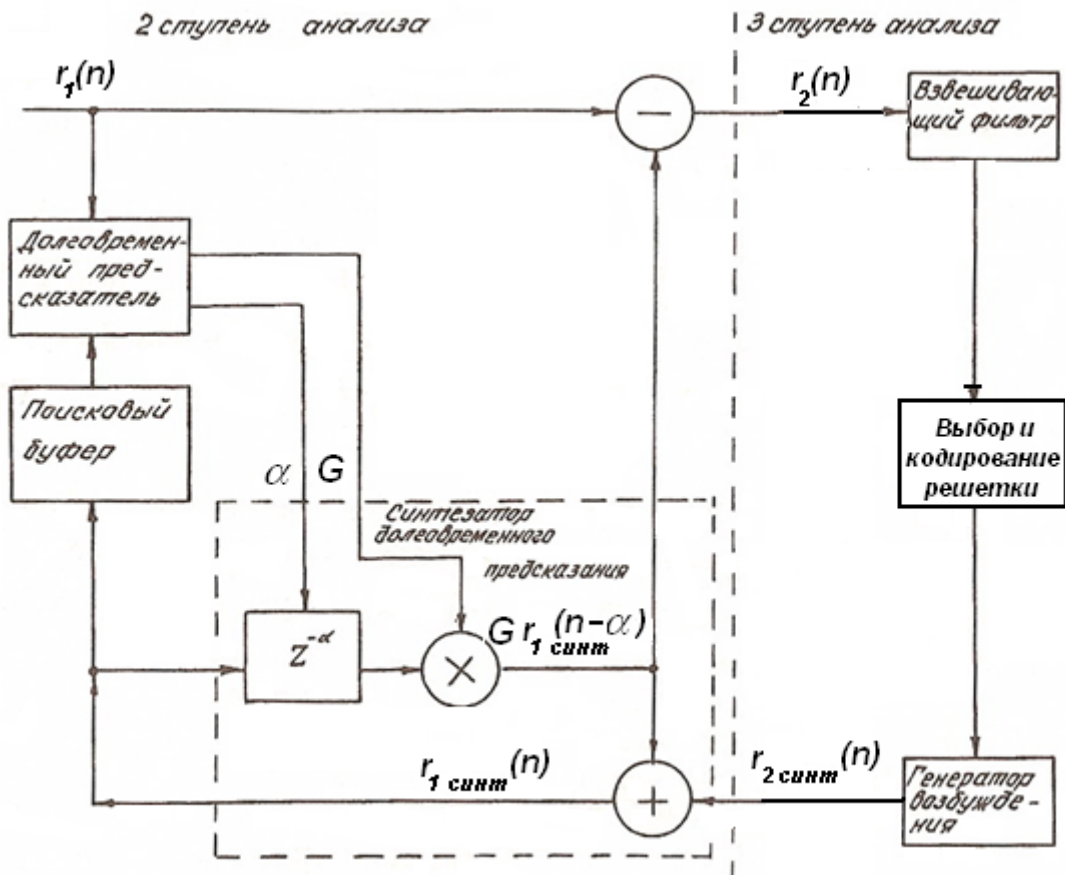


Рис.1.20. 2 и 3 ступени анализа в вокодере RPE-LTP

Значения G и α находят, как и в вокодере ACELP, минимизируя средне-квадратичное значение ошибки

$$E = \sum_{n=0}^{39} e_2^2(n) = \sum_{n=0}^{39} [r_1(n) - Gr_{1\text{синт}}(n - \alpha)]^2, \quad (1.48)$$

но их определяют и кодируют с меньшей точностью, чем в вокодере ACELP. Значения α (только целые) кодируют 7 битами, а G – 2 битами. Для передачи 4 пар $\{G; \alpha\}$ сегмента речи требуется 36 бит.

В вокодере RPE-LTP сигнал возбуждения $r_{2\text{синт}}(n)$ формируют, аппроксимируя второй остаточный сигнал $r_2(n)$, полученный после удаления основного тона (1.47). Для этого сигнал $r_2(n)$ каждого подсегмента прореживают, формируя решетку из 13 эквидистантных отсчетов. 27 остальных отсчетов подсегмента заменяют нулями. Для каждого подсегмента существуют 3 варианта решетки, начинающихся с 0, 1 или 2 отсчета. Оптимальную решетку выбирают по максимуму суммы квадратов 13 значимых отсчетов (максимуму мощности).

При передаче параметров решетки (сигнала $r_{2\text{синт}}(n)$) по радиоканалу для каждого подсегмента используют:

2 бита для кодирования смещения решетки,

6 бит для кодирования амплитуды максимального по величине отсчета в решетке,

3 бита для кодирования относительной амплитуды каждого из отсчетов решетки.

В результате необходимо 47 бит для передачи параметров одной решетки и 188 бит для 4-х решеток. Всего для передачи параметров 20-мс сегмента сжатой речи в вокодере RPE-LTP требуется $36+36+188=260$ бит, что соответствует информационной скорости 13 кбит/с.

В декодере, при синтезе речевого сигнала, вначале генерируют сигнал возбуждения $r_{2\text{синт}}(n)$ (решетку). Далее синтезируют основной тон

$$r_{1\text{синт}}(n) = r_{2\text{синт}}(n) + G r_{1\text{синт}}(n - a)$$

и форманты.

Первые вокодеры с кодовым возбуждением в стандартах сотовой связи появились в 1995г. в американском стандарте IS-95 (cdmaOne). В 1997-1999г. были опубликованы спецификации улучшенного полноскоростного вокодера (EFR – Enhanced Full Rate) для сетей GSM с информационной скоростью 12,2 кбит/с на основе структуры ACELP. Рассмотренный в 1.6 вокодер AMR был разработан в 2001г.

В 21 веке начали разрабатывать вокодеры для передачи широкополосного речевого сигнала (50 – 7000Гц) для сотовых сетей 3-го поколения. Первые спецификации вокодера AMR-WB (Adaptive Multi-Rate Wideband) появились в 2004г. После доработки в 2007г. выпущены спецификации Rel.7, без существенных изменений подтверждаемые до последнего Rel.11 [8].

Вокодер AMR-WB относится к классу ACELP и он построен на тех же принципах и технологиях, что и рассмотренный ранее в п.1.6 вокодер AMR. Как и в вокодере AMR, голосовой сигнал разбивают на 20мс сегменты, из которых при передаче удаляют форманты. Каждый сегмент делят на 4 подсегмента длительностью 5мс. Для каждого подсегмента находят оптимальный сигнал возбуждения и первый остаточный сигнал. Из-за расширения полосы передаваемых частот в 2 раза частота дискретизации исходного речевого сигнала в вокодере AMR-WB в 2 раза больше (16 кГц), чем в вокодере AMR (8 кГц). Каждая выборка кодирована 14-разрядным двоичным кодом. Однако обработку речи при ее сжатии в кодере ведут с пониженной “внутренней” частотой выборки 12,8 кГц, для чего исходный 20мс сегмент речи из 320 выборок подвергают децимации, формируя пакет из 256 выборок. В декодере, при синтезе речевого сигнала в приемнике, выполняют обратное преобразование.

Специфицированы 9 различных скоростей сжатой речи на выходе кодера AMR-WB: 6,6; 8,85; 12,65; 14,25; 15,85; 19,85; 23,05; 23,85 кбит/с. Номинальная скорость, обеспечивающая высокое качество речевого сигнала, 12,65 кбит/с. На более низкие скорости (8,85 и 6,6 кбит/с) переходят при работе в “плохих” каналах, с ухудшенным отношением сигнал/помеха, где снижают скорость избыточного кодирования. Скорости от 14,25 до 23,85 кбит/с предназначены для высококачественной передачи голоса и музыки по каналам сетей UMTS и GSM/EDGE, в том числе при передаче мультимедиа и услуг потокового трафика в пакетном режиме. Сейчас в выпускаемых мобильных телефонах используют 3 конфигурации скоростей вокодеров AMR-WB:

- конфигурация А (Config-WB-Code 0), скорости 6,6; 8,85; 12,65 кбит/с (обязательная конфигурация),

- конфигурация В (Config-WB-Code 2), скорости 6,6; 8,85; 12,65 и 15,85 кбит/с,
- конфигурация С (Config-WB-Code 4), скорости 6,6; 8,85; 12,65 и 23,85 кбит/с.

Коммерческий запуск передачи широкополосной речи начат операторами сетей 3G с 2010г.

Следующий по сложности широкополосный **кодек AMR-WB+** (Extended Adaptive Multi-Rate – Wideband) дорабатывался в 2004 – 2007гг; спецификации в конечном виде выпущены в Rel.8 в 2008г. Кодек предназначен прежде всего для сжатия голоса, музыки и прочего звукового сопровождения при передаче потокового видео и различных вариантов мультимедийного трафика. При этом возможна передача как моно, так и стереовещания [9].

Кодек AMR-WB+ может работать как.

AMR-WB с частотой выборок исходного сигнала 16 кГц (14 бит на выборку) и с внутренней частотой при сжатии сигнала 12,8 кГц. AMR-WB+ поддерживает также частоты выборок 16/24/32/48 кГц при кодировании каждой выборки 16-битовым кодом. Возможности кодека AMR-WB+ значительно шире, чем у вокодера AMR-WB. Кодек AMR-WB+ может сжимать и синтезировать как сигналы моно вещания, так и стерео вещания.

На входе кодека AMR-WB+ передаваемый сигнал разбивают на кадры. Каждый кадр состоит из 2048 выборок преобразованного входного сигнала, следующих с внутренней частотой обработки в кодере $f_s = 12,8...38,4$ кГц. Эти 2048 отсчетов в свою очередь делят на 2 суперкадра по 1024 выборки в каждом: один суперкадр для низкочастотной части сигнала НЧ, другой – для высокочастотной ВЧ. НЧ-сигнал включает в себя составляющие, расположенные ниже частоты $f_s/4$. Компоненты с частотами выше $f_s/4$ попадают в ВЧ-сигнал. Далее суперкадр может быть фрагментирован на 4 кадра по 256 выборок в каждом. Выборки с внутренней частотой f_s получают с помощью специального алгоритма преобразования частоты выборок входного сигнала.

Для сжатия низкочастотной части сигнала НЧ и высокочастотной части ВЧ используют разные технологии. Сигнал НЧ обрабатывают по рассмотренному ранее протоколу ACELP или по протоколу кодового возбуждения с преобразованием TCX (Transform Coded Excitation). В протоколе TCX при сжатии аудио сигнала и определении параметров для его синтеза соответствующие операции производят не над самим сигналом (его выборками), а над отсчетами спектра, полученного после Фурье-преобразования кадров из 256, 512 или 1024 выборок. Протокол ACELP дает лучшие результаты при передаче речи, а протокол TCX – музыки. В кодеке AMR-WB+ использован адаптивный метод выбора протокола (мода) обработки сигнала. Кодер параллельно обрабатывает сигнал, разбитый на кадры, по ACELP, TCX протоколам и выбирает вариант, минимизирующий относительную среднеквадратичную ошибку за кадр при синтезе сигнала. При этом для варианта TCX просматривают кадры разной длины: из 256, 512 и 1024 выборок.

Сигнал ВЧ кодируют с довольно низкой скоростью, применяя технологию расширения полосы BWE (Bandwidth Extention). Суть метода BWE состоит в том, что при синтезе ВЧ-сигнала используют сигнал возбуждения, полученный для НЧ-сигнала, но при синтезе формант оперируют коэффициентами линейного предсказателя \bar{a}_k (1.26) ВЧ-сигнала.

На рис.1.21. представлена структура передающей части кодера AMR-WB+ при 2-х возможных вариантах входного сигнала: моно (М) и стерео (П – правый, Л – левый). При передаче моно сигнала его разделяют на низкочастотную $M_{нч}$ и высокочастотную $M_{вч}$ части. Сигнал $M_{нч}$ кодируют по протоколу ACELP или по протоколу TCX; при кодировании высокочастотной части используют метод BWE.

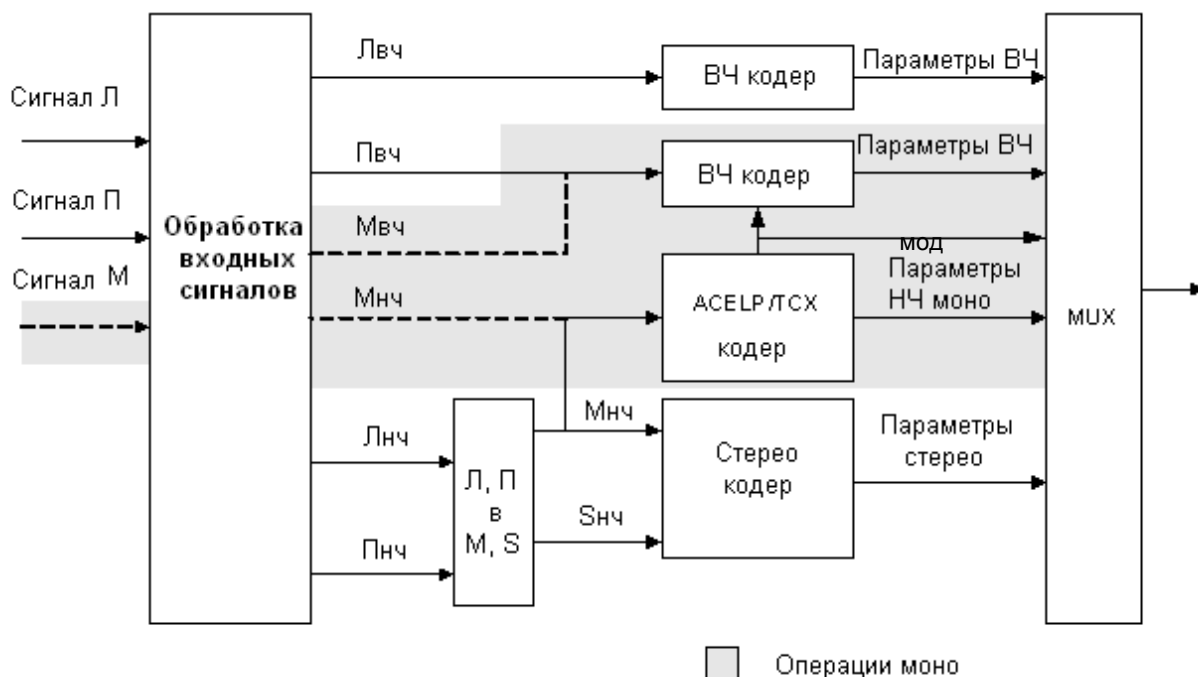


Рис.1.21. Структура кодера AMR-WB+

При стерео передаче сигналы левого и правого каналов разделяют на НЧ и ВЧ части. Далее $P_{нч}$ и $L_{нч}$ объединяют в единый моно сигнал $M_{нч}$, который кодируют по протоколу ACELP или TCX. Кроме моно НЧ-параметров, передают дополнительные стерео-параметры, необходимые для синтеза стерео сигнала в приемнике. Кодирование ВЧ-сигналов осуществляют по методу BWE.

В декодере AMR-WB+ (рис. 1.22) происходит синтез ВЧ и НЧ компонент сигналов с их последующим объединением по соответствующим каналам.

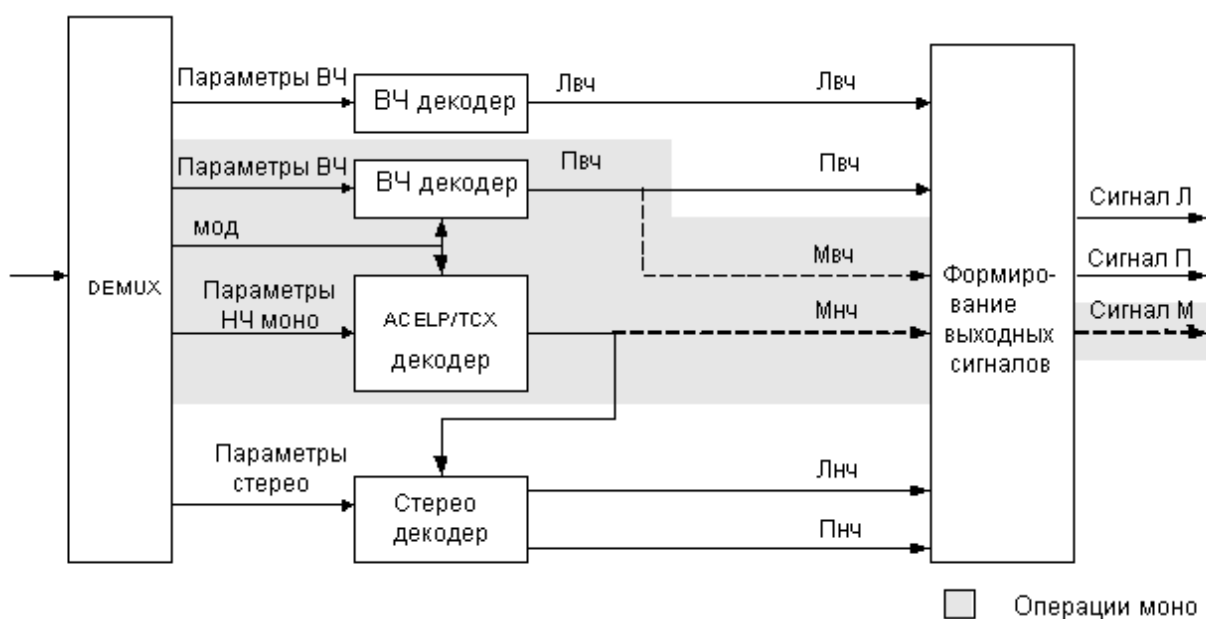


Рис.1.22. Структура декодера AMR-WB+

В аппаратуре стандарта CDMA2000 используют вокодеры семейства 4GV (Qualcomm's Fourth generation Vocoder) с кодовым возбуждением, куда входят вокодеры EVRS-B (Enhanced Variable Rate Codec) для передачи узкополосной речи и EVRS-WB для передачи широкополосной речи. В вокодерах EVRS-B телефонию передают со скоростями 8,55 кбит/с и 4 кбит/с. В широкополосных вокодерах информационный сигнал разделяют на два: НЧ (50 – 4000Гц) и ВЧ (3500 – 7000), которые обрабатывают и передают независимо.

1.8. Прерывистая передача речи

Во всех цифровых стандартах сотовой связи используют технологию прерывистой передачи речи DTX (Discontinuous Transmission) по выделенному каналу. В сетях стандарта GSM в паузах речи передатчики соответствующей базовой и абонентской станции молчат, а в сетях стандарта UMTS прерывается передача по выделенному каналу данных DPDCH и передают только сопровождающий канал управления DPCCCH [10]. В результате снижается уровень помех, освобождаются каналы на радиоинтерфейсе, в транспортных сетях и увеличивается время работы аккумуляторной батареи мобильного терминала.

Для реализации DTX при телефонии используют детекторы активности речи VAD (Voice Activity Detector). Задача VAD состоит в определении активной фазы речи и в передаче во время пауз короткого пакета "комфортного шума" для обеспечения естественности восприятия передаваемого сигнала. Для разных типов вокодеров разработаны свои VAD. Ниже рассмотрены структура и основные процедуры функционирования VAD для вокодера AMR.

Блок-схемы обработки сигнала телефонии при передаче и приеме приведены на рис. 1.23 [11]. При передаче со стороны подсистемы радиодоступа RAN (Radio Access Network) из 8-битового телефонного потока из сети (1), компрессированного по А-закону, восстанавливают 13-битовый сигнал ИКМ (2). В мобильном терминале MS/UE¹ аналоговый сигнал с микрофона после ФНЧ в АЦП также преобразуют в 13-битовый ИКМ-сигнал.

Анализируя параметры поступающих 20-мс кадров (сегментов) информационного потока, VAD устанавливает флаг наличия речевого (или другого) сигнала (3) и определяет тип передачи: речь (активная фаза) или шум (6). В активной фазе вокодер передает (7) информационные кадры для синтеза сигнала в приемнике, а во время пауз – кадры SID (Silence Description) для синтеза комфортного шума (5).

На приемную сторону поступают информационные кадры (8) или кадры SID, а также, а также сообщение о типе полученного кадра и индикации поврежденных кадров BFI (Bad Frame Indication). При установке индикатора BFI полученный кадр отбрасывают и заменяют последним принятым кадром. Такая замена информационных кадров позволяет убрать разрывы в передаче речи на интервале до 120 мс.

¹ MS (Mobile Station) – терминал стандарта GSM; UE (User Equipment) – терминал стандарта UMTS.

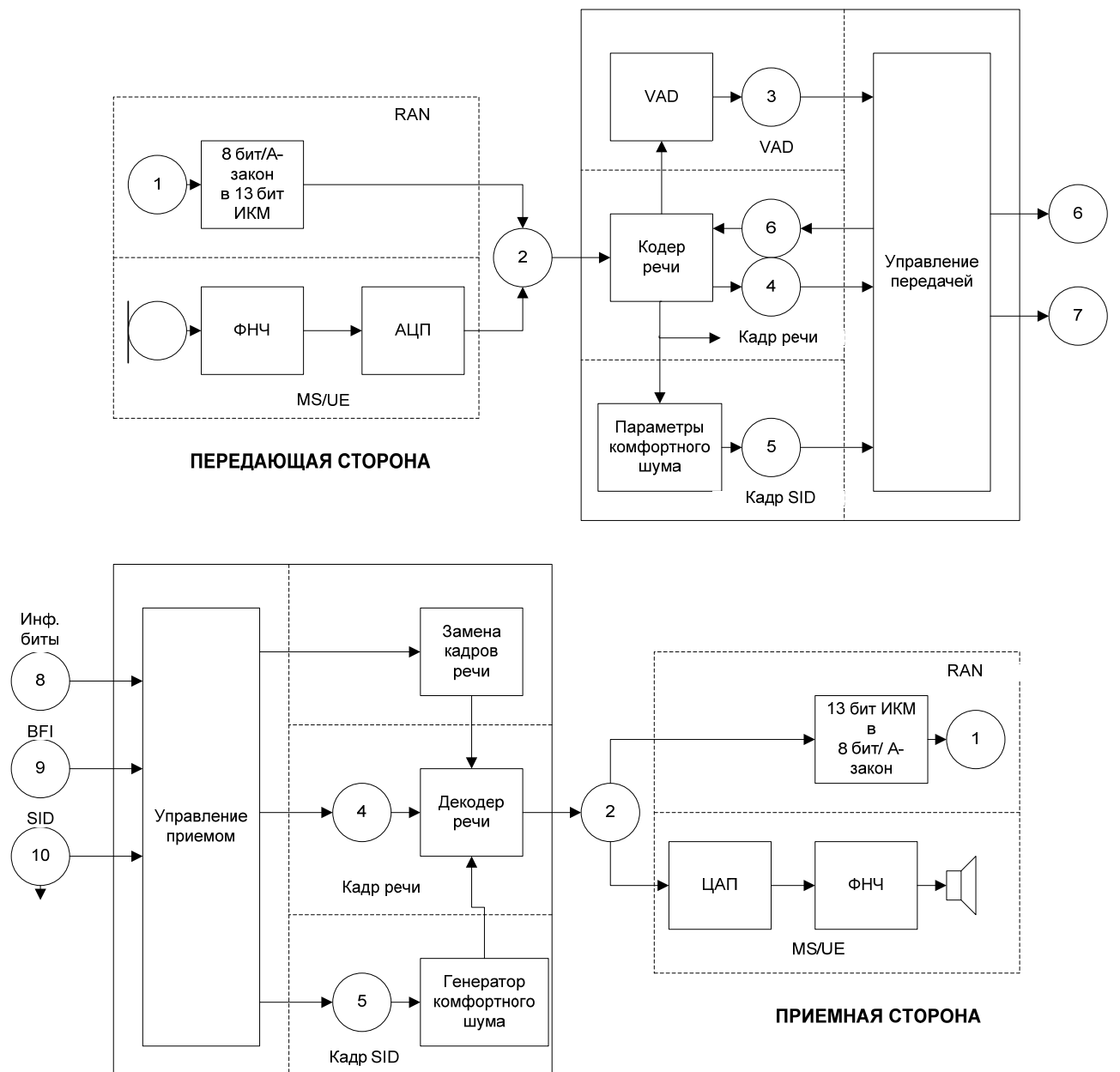


Рис. 1.23. Блок-схема обработки сигнала телефонии

В GSM/UMTS специфицировано 2 алгоритма VAD [12]. Первый, более поздний алгоритм, ориентирован на кодеки типа CELP (ACELP), второй, используемый с 1991г., описан в [2]. Рассмотрим, как работает VAD по первому алгоритму, который проиллюстрирован рис. 1.24.

Входной сигнал $s(n)$, разделенный на 20мс сегменты, пропускают через 9 полосовых фильтров (*Filter bank*) и измеряют мощность сигналов на выходах всех фильтров ($level[n]$). Полосы пропускания фильтров приведены в табл. 1.4.

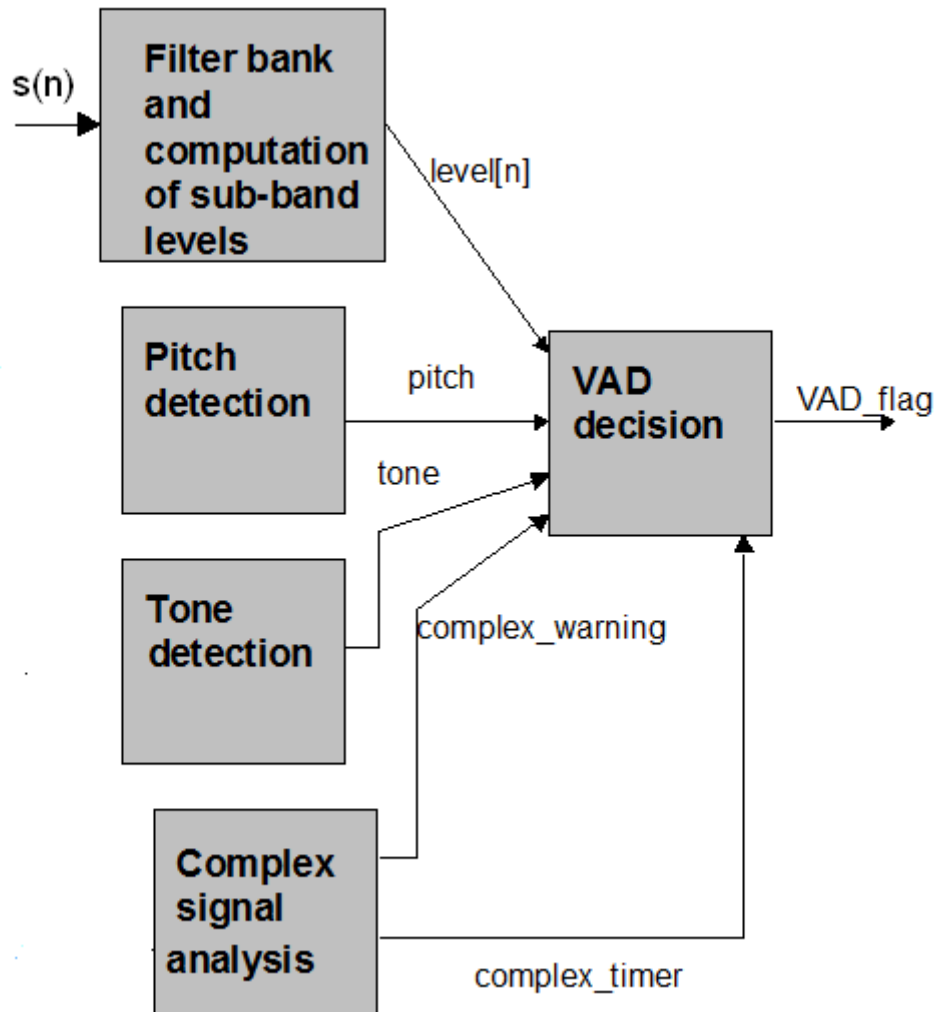


Рис. 1.24. Упрощенная блок-схема алгоритма VAD, вариант 1

Таблица 1.4

Номер полосы	Граничные частоты, Гц
1	0 – 250
2	250 – 500
3	500 – 750
4	750 – 1000
5	1000 – 1500
6	1500 – 2000
7	2000 – 2500
8	2500 – 3000
9	3000 – 4000

Далее определяют наличие основного тона (*Pitch detection*) в открытой петле при определении параметров первого остаточного сигнала (см. п.1.6).

Tone detection обеспечивает индикацию отдельных тонов, о наличии которых выставляют флаг *tone*. Анализатор комплексного сигнала позволяет обнаружить сложные сигналы, например, музыку. Здесь выявляют корреляционные зависимости прежде всего в верхней части частотного диапазона. При их обнаружении выставляют сообщение *complex_warning*.

Главная задача VAD – определить, что передают в текущем сегменте: речевой сигнал, отдельные тоны, музыку или шум. Для этого в паузах передачи речи VAD оценивает параметры *фонового шума*. Паузу в речевом сигнале VAD определяет по мощности сигнала на входе вокодера и по отсутствию основного тона (процедура *Pitch detection*). При этом мощность фонового шума находят, суммируя мощности сигналов на выходах всех полосовых фильтров:

$$noise_level = \sum_{n=1}^9 bckr_est(n), \quad (1.49)$$

где $bckr_est(n)$ – уровень фонового шума на выходе одного фильтра.

Конец паузы и начало активной фазы речи определяют по превышению уровня суммарного сигнала (речь и шум) над уровнем фонового шума. Эту разницу оценивают по отношению

$$snr_sum = \sum_{n=1}^9 MAX\left(1, 0; \frac{level(n)}{bcbk_est(n)}\right)^2 \quad (1.50)$$

В (1.50) $level(n)$ – уровень суммарного сигнала на выходе соответствующего фильтра n . Отношение snr_sum должно превышать шумовой порог, зависящий от уровня фонового шума. Чем он выше, тем ниже шумовой порог. Это позволяет выявлять и обрабатывать слабые речевые сигналы на фоне сильных шумов, хотя при этом и увеличивается вероятность обработки сильного шума вместо речи.

Второй признак, обязательный для речевого сигнала – обнаружение в нем основного тона, выполняемое в блоке *Pitch detection*. В блоках *Pitch detection*, *Tone detection* и *Complex signal analysis* исследуют взвешенные входные сигналы $s_w(n)$. Как известно из алгоритма ACELP, в открытой петле определяют ориентировочную задержку α фильтра долговременного предсказателя. Если в процессе поиска α не удастся найти требуемой корреляционной зависимости в течение двух последовательных 20 мс сегментов входного сигнала, то устанавливается *pitch flag*, а VAD приступает к дополнительному исследованию содержания пришедшего сигнала. VAD выявляет, не содержит ли передаваемый сигнал отдельных тонов (*Tone detection*) или сложных периодических сигналов (*Complex signal analysis*). Замена этих сигналов комфортным шумом ухудшает естественность передачи речи.

Во время пауз речи абонент слышит усредненный фоновый шум, называемый комфортным шумом. Для его генерации берут данные из следующих подряд 8 сегментов входного сигнала общей длительности 160 мс [13]. Анализируя этот фрагмент шума, VAD определяет его спектральные характеристики, которые представляют в виде 10 составляющих вектора f^{mean} линейных спектральных пар. В памяти вокодера записаны 8 характерных вариантов векторов фонового шума f^{ref} . VAD вычисляет ошибку

$$e(i) = f^{mean}(i) - f^{ref}(i) \quad (1.51)$$

и выбирает вектор по минимуму $\sum_{i=1}^{10} e^2(i)$.

Далее VAD формирует кадр SID. В него входит номер вектора f^{ref} (3 бита), 26 бит для передачи квантованной ошибки $e(i)$ в форме соответствующих спектральных пар. Наконец, 6 бит используют для передачи логарифма мощности входного сигнала, подвергнутого предварительной высокочастотной фильтрации.

Таким образом, кадр SID содержит 35 информационных бит. Эти кадры во время пауз в речи передают каждые 480 мс в одном коротком кадре. Остальное время по выделенному каналу соответствующий передатчик молчит.

Контрольные вопросы к главе 1

1. Как осуществляют преобразование аналогового телефонного сигнала в цифровой.
2. Зачем производят компрессию цифрового речевого сигнала.
3. В чем состоит алгоритм дифференциальной ИКМ. За счет чего происходит снижение скорости передачи речи.
4. В чем состоит алгоритм дельта-модуляции. За счет чего происходит снижение скорости передачи речи.
5. Опишите процедуру обработки речи в вокодере.
6. Почему вокодеры строят по схеме анализ-синтез.
7. Какие элементы входят в функциональную схему вокодера AMR.
8. Какую роль выполняет в вокодеру AMR взвешивающий фильтр.
9. Как в вокодеру AMR определяют параметры формант.
10. Какие параметры характеризуют основной тон.
11. Как в вокодеру AMR определяют параметры основного тона.
12. Что означает "кодовое возбуждение" в вокодерах. По каким критериям ищут оптимальный код возбуждения.
13. Как реализована алгебраическая кодовая книга в вокодеру AMR.
14. Опишите алгоритм поиска кода возбуждения в вокодеру AMR.
15. Как построен декодер вокодера AMR.
16. Опишите структуру вокодера RPE-LTP.
17. Укажите отличия в структуре и алгоритмах работы вокодеров AMR и RPE-LTP.
18. В чем отличия вокодеров AMR-WB и AMR-WB+ от вокодера AMR.
19. Какие принципы положены в основу работы вокодера AMR-WB.
20. Как происходит сжатие и синтез сигналов в вокодеру AMR-WB+.
21. Зачем в сотовых сетях используют прерывистую передачу речи.
22. Опишите алгоритм работы детектора активности речи VAD.
23. Как в VAD отличают речевой сигнал от фонового шума.
24. Как формируют и синтезируют в вокодеру AMR фоновый шум.

2. КОДИРОВАНИЕ ИНФОРМАЦИОННЫХ СИГНАЛОВ

2.1. Необходимость кодирования информационных сигналов

Во всех стандартах сотовой связи используют избыточное кодирование информационных сигналов при их передаче по радиоканалам. Это обусловлено нестабильностью параметров каналов в результате быстрых и медленных замираний, что может приводить к появлению ошибок при приеме. Для индикации ошибок и их исправления и используют защитное избыточное кодирование.

Обнаружение ошибок осуществляют с помощью достаточно простых блочных кодов. Следующая, более важная и сложная ступень задачи кодирования – выявление ошибок и их исправление. Для ее выполнения в системах радиодоступа широко применяют сверточные коды. Использование сверточных кодов требует наличия мощных быстродействующих процессоров. Сверточные коды хорошо исправляют случайные единичные, двоянные или строенные ошибки, но бесполезны при пакетах ошибок. Поэтому сверточное кодирование совмещают с перемежением информационных символов.

Блочные коды, исправляющие ошибки, также требуют при декодировании мощных вычислительных средств. При этом такие коды, как коды Файра и Рида-Соломона, способны исправлять пакеты ошибок. При передаче сигнализации часто используют каскадное кодирование. К внутреннему сверточному кодированию добавляют внешнее с исправлением пакетов ошибок. Разумеется, такое сложное кодирование требует большой избыточности, что формально снижает эффективность использования каналов, но существенно увеличивает достоверность принимаемой информации. Избыточность передаваемой информации определяют по скорости кода: $R_{\text{кода}} = n_{\text{инф}}/n_{\text{код}}$, где $n_{\text{инф}}$ – число кодируемых информационных бит, а $n_{\text{код}}$ – число бит после кодирования. Чем ниже скорость кода, тем большая кодовая защита принята в канале, но тем меньше скорость передачи информации при фиксированной скорости передачи в канале. При $R_{\text{кода}} = 1$ кодовая защита отсутствует.

В публикациях по обработке информационных и управляющих сигналов в системах подвижной связи часто говорят о кодировании, употребляя термин "предкоррекции" передаваемой информации". В традиционной радиотехнике под предкоррекцией понимают предискажения, компенсирующие искажения в канале передачи. Применительно к задачам кодирования речь идет о статистической предкоррекции сигнала, когда известны лишь статистические свойства канала, вносящего ошибки. Кодированию подвергают типовые единицы передачи информации: кадры или блоки. Методы и алгоритмы избыточного кодирования описаны в [4], [14], [15], [16].

2.2. Блочное кодирование

В блочных кодах (БК) кодирование информации производят поблочно, в группах определенного числа бит (блоке). При этом во многих случаях к k информационным символам добавляют проверочные символы, определяющие избыточность кода. Длина кодового слова с символами паритета составляет n символов. Такой метод кодирования называют систематическим, а соответствующие ему БК систематическими.

где

$$\mathbf{i} = \mathbf{i} \times \mathbf{I},$$

$$\mathbf{I} = \begin{vmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 1 \end{vmatrix} \quad (2.2)$$

единичная диагональная матрица k -го порядка, и объединяя (суммируя) левые и правые части (2.1) и (2.2), получаем матричную форму записи кодового слова систематического кода:

$$\mathbf{C} = \mathbf{i} + \mathbf{b} = \mathbf{i} \times \mathbf{G},$$

где $\mathbf{G} = [\mathbf{I} \ \mathbf{P}]$ называют порождающей матрицей систематического кода. Она имеет размер $[k \times n]$.

При декодировании принятого кодового слова используют проверочную матрицу \mathbf{H} размера $[(n-k) \times n]$, удовлетворяющую условию ортогональности,

$$\mathbf{G} \times \mathbf{H}^T = \mathbf{0}, \quad (2.3)$$

где T – символ транспонирования.

Если при прохождении по каналу связи в кодовом слове \mathbf{c} появятся ошибки, так что принятое слово $\mathbf{W} = \mathbf{C} + \mathbf{e}$, где \mathbf{e} – вектор ошибок, то, выполняя при декодировании операцию

$$\mathbf{W} \times \mathbf{H}^T = (\mathbf{C} + \mathbf{e}) \times \mathbf{H}^T = \mathbf{i} \times \mathbf{G} \times \mathbf{H}^T + \mathbf{e} \times \mathbf{H}^T = \mathbf{e} \times \mathbf{H}^T, \quad (2.4)$$

получаем вектор синдрома $\mathbf{S} = \mathbf{W} \times \mathbf{H}^T = \mathbf{e} \times \mathbf{H}^T$ длиной $(n-k)$, зависящий только от вектора ошибок и проверочной матрицы. Если ошибки отсутствуют, то $\mathbf{S} = \mathbf{0}$.

Для блочных кодов разработано большое число алгоритмов декодирования для поиска и исправления ошибок. Наиболее простой, но самый затратный метод – сравнить принятый вектор \mathbf{w} со всеми возможными векторами \mathbf{c} , соответствующим информационным словам \mathbf{i} . Тот вектор, который по Хэммингу минимально отличается от принятого вектора \mathbf{w} , считаем наиболее достоверным результатом.

Процесс поиска можно пояснить с помощью двумерной модели на плоскости (рис. 2.1). Каждое кодовое слово – точка в n -мерном пространстве. Минимальное расстояние по Хэммингу d_{min} определяют как минимум расстояния по всем возможным парам различных слов.

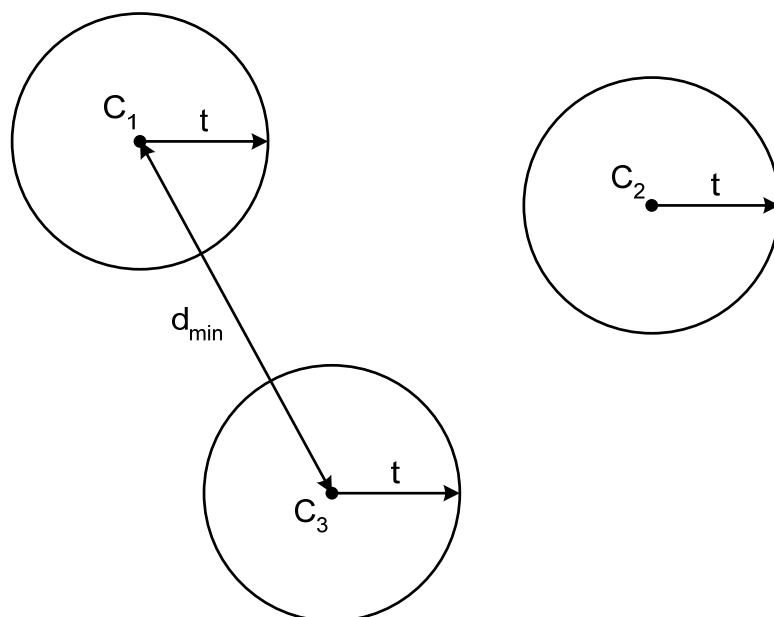


Рис. 2.1. К определению корректирующей способности кода

Корректирующая способность кода – это сфера, определяемая числом бит, которое может быть исправлено (радиус t на рис. 2.1). Общее число обнаруженных e_d и исправленных e_c ошибок подчиняется правилу

$$e_d + e_c \leq d_{min} - 1, \quad \text{причем } e_c \leq e_d.$$

Так, если кодовое расстояние $d_{min} = 7$, то можно обнаружить и исправить 3 ошибки (радиус $t = 3$). Можно построить код так, что будут обнаружены 4 ошибки, но исправлены только 2 (сузить радиус t). Во всяком случае, если в принятом слове будет более 4-х ошибок, то будет принят неверный результат, т.е. **процесс декодирования – всегда вероятностный процесс**.

Возможность применения регулярных алгоритмов декодирования обусловили широкое использование циклических БК. Если код циклический, то в кодовом слове $C = |C_{n-1}, C_{n-2} \dots C_1, C_0|$ можно сдвигать символы циклически на j позиций. При этом новое кодовое слово $C^{(j)} = |C_{n-1-j}, C_{n-2-j} \dots C_{n+1-j}, C_{n-j}|$ будет также разрешенным кодовым словом. При анализе циклических кодов C , b , i и G представляют в виде **полиномов**. Так, блоку двоичных чисел 10101111 соответствует информационный полином $(k-1)$ порядка $i(z) = z^7 + z^5 + z^3 + z^2 + z + 1$. При систематическом кодировании информационный полином сдвигают влево на $(n-k)$ разрядов и добавляют проверочный код, состоящий также из $(n-k)$ разрядов. Этим операциям соответствует алгебраическая запись

$$C(z) = i(z) \times z^{n-k} + b(z) \quad (2.5)$$

Следовательно, многочлен $C(z)$ является полиномом $(n-1)$ порядка

$$C(z) = C_{n-1} z^{n-1} + C_{n-2} z^{n-2} + \dots + C_1 z^1 + C_0, \quad (2.6)$$

где для двоичных кодов в $GF(2)$ коэффициенты $C_m = 0$ или 1.

В циклических кодах кодовый полином $C(z)$ делится без остатка на порождающий полином $G(z)$:

$$R_{G(z)}[C(z)] = 0 \quad (2.7)$$

Подстановка (2.5) в (2.7) позволяет рассчитать проверочную часть $b(z)$.

$$R_{G(z)}[i(z) \times z^{n-k}] + R_{G(z)}[b(z)] = 0,$$

откуда

$$b(z) = -R_{G(z)}[i(z) \times z^{n-k}], \quad (2.8)$$

так как степень полинома $b(z)$ меньше степени полинома $G(z)$.

Синдром в циклических кодах также определяют алгебраически в виде полинома $S(z)$. По аналогии с рассмотренным ранее векторным представлением запишем принятое слово как $W(z) = C(z) + e(z)$, где $e(z)$ – ошибка, внесенная каналом связи. Используя (2.7), получаем

$$S(z) = R_{G(z)}[W(z)] = R_{G(z)}[C(z) + e(z)] = R_{G(z)}[e(z)] \quad (2.9)$$

Когда $S(z) = 0$, ошибки в принятом блоке отсутствуют. Если они есть, то в синдроме $S(z)$ появятся члены, отличные от нуля. Для примера найдем синдром для кода Хэмминга (7,4,1), исправляющего одиночные ошибки. Кодовое слово $C(z)$ содержит 7 бит, из которых старшие 4 составляют передаваемый информационный блок. Порождающий многочлен $G(z) = z^3 + z + 1$. Положим, что ошибка содержится в бите z^5 . Это значит, что $e(z) = z^5$ (напомним, что речь идет о двоичном коде). Процесс деления z^5 на $G(z)$ в $GF(2)$ представлен ниже:

$$\begin{array}{r} z^5 \\ \underline{z^5 + z^3 + z^2} \\ z^3 + z + 1 \\ \underline{z^3 + z + 1} \\ z^2 + z + 1 \end{array}$$

где остаток $R_{G(z)}[e(z)] = z^2 + z + 1$ и есть искомым синдром $S(z)$.

Циклические коды широко используют при расчете контрольных сумм для обнаружении ошибок при приеме информационных блоков (кадров). Так, в стандарте UMTS при формировании кода паритета блоков транспортных каналов в зависимости от размера блока используют CRC (Cycle Redundancy Code) длиной в 24, 16, 12 или 8 бит. Этим $CRC = b(z)$ соответствуют порождающие полиномы [17]:

$$\begin{aligned} G_{24}(z) &= z^{24} + z^{23} + z^6 + z^5 + z + 1 \\ G_{16}(z) &= z^{16} + z^{12} + z^5 + 1 \\ G_{12}(z) &= z^{12} + z^{11} + z^3 + z^2 + z + 1 \\ G_8(z) &= z^8 + z^7 + z^4 + z^3 + z + 1 \end{aligned}$$

Покажем на примере, как рассчитать CRC при использовании порождающего полинома $G_8(z)$. Пусть передаваемый блок представляет собой последовательность нулей, заканчивающихся словом 000010³ (рис. 2.2). Этому слову соответствует полином

$$i(z) = z. \quad (2.10)$$

Кодовое слово $C(z)$ определим согласно (2.5), где $b(z)$ надо получить и передать по каналу связи после $i(z)$. Вычислим $b(z)$ с помощью (2.8):

$$b(z) = -R_{G(z)}[i(z) \times z^{n-k}] = -R_{G(z)}(z^9).$$

Процесс деления z^9 на $z^8 + z^7 + z^4 + z^3 + z + 1$ приведем пооперационно:

³ Реальная длина блоков, передаваемых по радиоканалам, составляет сотни и тысячи бит.

$$z^9 \quad / (z^8 + z^7 + z^4 + z^3 + z + 1) = z + 1$$

$$\frac{z^9 + z^8 + z^5 + z^4 + z^2 + z}{z^8 + z^5 + z^4 + z^2 + z}$$

$$\frac{z^8 + z^7 + z^4 + z^3 + z + 1}{z^7 + z^5 + z^3 + z^2 + 1}$$

Остаток от деления $b(z) = -(z^7 + z^5 + z^3 + z^2 + 1) = z^7 + z^5 + z^3 + z^2 + 1$ (в GF(2)!) и есть искомая контрольная сумма. Рассчитанная при приеме блока, она должна сойтись с переданной. Полное кодовое слово $C(z) = i(z) \times z^{15-7} + b(z)$ показано на рис. 2.2.

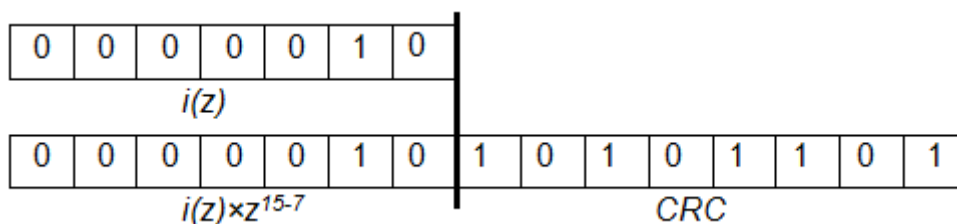


Рис. 2.2. Формирование блока с контрольной суммой

2.3. Сверточное кодирование.

Сверточное кодирование – основной способ избыточного кодирования данных в сетях радиодоступа. В сверточных кодах используют непрерывную последовательную обработку потока информационных символов. Кодер обладает памятью в том смысле, что символы на его выходе зависят не только от очередного фрагмента (символа) на его входе, но и от предыдущих символов.

Свёрточный код (СК) относится к числу непрерывных кодов. В общем случае при сверточном кодировании группу из k информационных бит в канале передачи заменяют на n выходных двоичных бит, при формировании которых используют $(K-1)$ предшествующих групп информационных бит. Величину K называют длиной кодового ограничения. Таким образом, свёрточный код характеризуют 3 параметра (n, k, K) . В системах подвижной связи широко используют СК, где $k=1$, $n=2$ или 3 (одному информационному биту соответствует 2 или 3 передаваемых бита после кодирования). Часто встречаются варианты, когда $k=2$, $n=3$ (двум информационным битам соответствуют 3 бита кодированного сигнала). Отношение $R_{\text{кода}} = k/n$, характеризующее вносимую СК избыточность, определяет скорость кодирования. Для приведенных примеров СК скорости $R_{\text{кода}} = 1/2$, $R_{\text{кода}} = 1/3$ и $R_{\text{кода}} = 2/3$ соответственно. Большие скорости кода позволяют увеличить пропускную способность канала связи, однако снижение скорости $R_{\text{кода}}$ улучшает качество передачи, уменьшает коэффициент ошибок (BER) на выходе приемника. Произведение $k \times K$ равно числу информационных бит, участвующих в формировании n выходных бит СК.

Закон кодирования определяет n образующих полиномов $q_1 \dots q_n$. Для простейшего СК (2,1,3)

$$q_1 = 1 + z^2 \quad (2.10)$$

$$q_2 = 1 + z + z^2 \quad (2.11)$$

где $+$ обозначает операцию сложения по модулю 2, а z^i – оператор задержки на i бит.⁴

⁴ В ряде публикаций, например в [4], при записи образующих полиномов используют отрицательные степени z , поскольку в теории z-преобразований задержка на m тактов соответствует z^{-m} .

Рассмотрим как пример процесс СК последовательности из 5 бит $c_i = 01101$ по закону образующих полиномов (2.10) и (2.11). Каждому информационному биту C_i соответствуют 2 выходных бита: q_1 и q_2 , в формировании которых участвуют 2 предшествующих информационных бита – C_{i-1} , C_{i-2} ($K=3$). При СК первого бита считаем, что в "предшествующие" биты записаны логические нули. Результат СК показан на рис.2.3. При кодировании каждого следующего бита информационный сигнал сдвигают на 1 бит вправо и повторяют процесс.

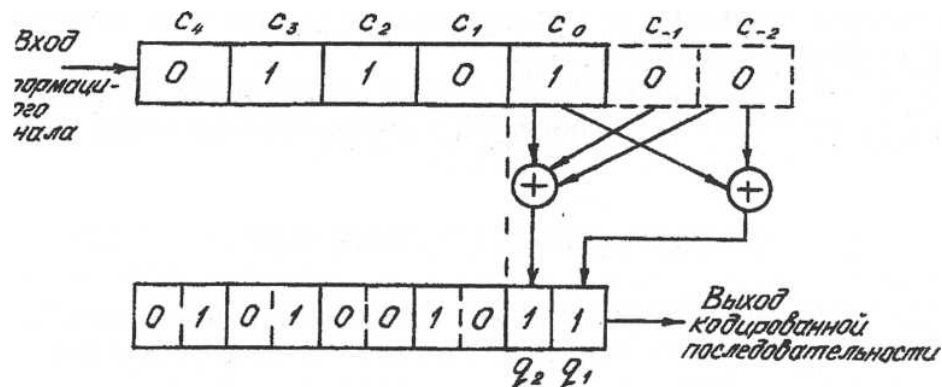


Рис. 2.3. Сверточный кодер (2,1,3)

Из рис.2.3 следует, что последние входные биты: C_3 и C_4 , участвуют не в 3 кодовых комбинациях, а в двух и одной соответственно. Чтобы обеспечить одинаковую кодовую защиту всех информационных бит, в конце каждого информационного блока добавляют $K-1$ хвостовых бит (обычно логические нули). При размерах информационных блоков в десятки, сотни и тысячи бит добавление хвостовых бит практически не сказывается на скорости кодирования.

Алгоритмы приема (декодирования) сигнала основаны на оценке его состояний в тактовые моменты времени. Под состоянием сигнала будем понимать отрезок информационного сигнала, состоящий из $(K-1) \times k$ информационных бит.

При $k=1$ каждое состояние сигнала определяют по $(K-1)$ последним принятым битам. Всего при $k=1$ сигнал может принимать 2^{K-1} состояний. В рассматриваемом примере СК (2,1,3) возможны 4 состояния сигнала: 00, 10, 01 и 11. Переход из одного состояния в другое при приеме очередного бита удобно представлять на решетчатой диаграмме состояний сигнала (рис.2.4). Из каждого состояния возможны 2 перехода в другие состояния в зависимости от того, пришла логическая 1 или 0. Каждому пути (переходу) соответствует своя комбинация кодовых бит согласно (2.10) и (2.11) (рис.2.5).

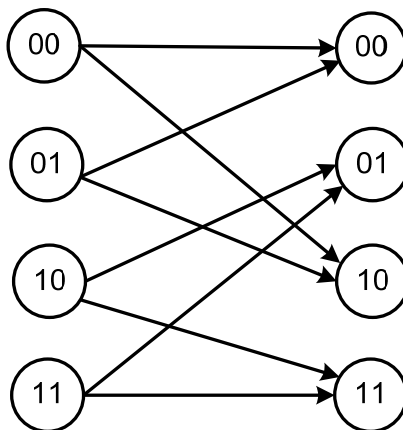


Рис. 2.4. Диаграмма изменения состояний сигнала в кодере (2,1,3)

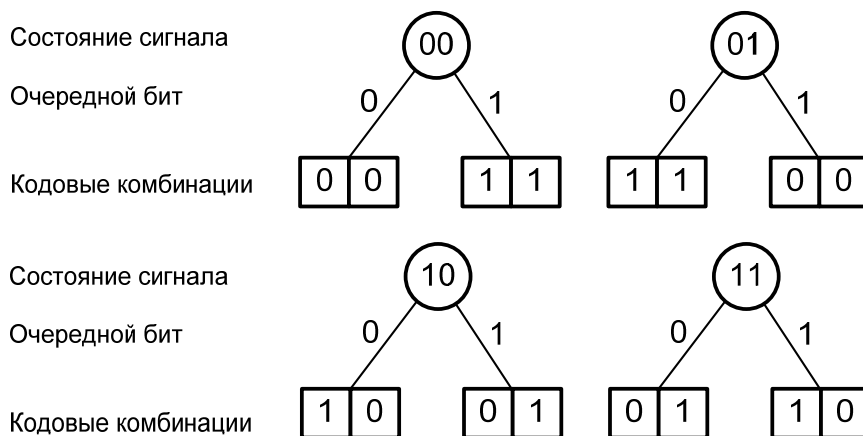


Рис. 2.5. Возможные переходы в кодере (2,1,3)

Процесс декодирования сигнала, подвергнутого СК, состоит и определении оптимального пути из одного состояния к другому в соответствии с алгоритмом кодирования (полиномами $q_1 \dots q_n$). Под оптимальным путем понимают путь, удовлетворяющий критерию максимального правдоподобия. Сегодня наиболее эффективным алгоритмом, позволяющим найти оптимальный путь, считают алгоритм Витерби, названный так по имени предложившего его А. Дж. Витерби. Идея алгоритма Витерби состоит в том, что в декодере воспроизводят все возможные пути последовательных изменений состояний сигнала, сравнивая при этом правильные кодовые комбинации, соответствующие законам кодирования $q_1 \dots q_n$ с принятыми комбинациями по каналу связи. Анализируя ошибки между принятыми битами и требуемыми комбинациями, определяют оптимальный путь.

Продемонстрируем работу алгоритма Витерби для рассматриваемого примера СК (2,1,3). Возьмем последовательность символов, полученную на рис.2.3, и будем полагать, что все биты q_i приняты без ошибок (рис.2.6). Пояснения сопровождается графическое отображение изменений состояния сигнала на диаграмме (рис.2.7). Начальным всегда является состояние 00. В тактовый момент $z=1$ возможны два перехода: $00 \rightarrow 00$, $00 \rightarrow 10$. Первому переходу соответствует кодовая комбинация $q_2q_1 = 00$, второму – $q_2q_1 = 11$ (см. рис.2.5). Сравнение с принятыми q_2q_1 показывает, что первой ветви $00 \rightarrow 00$ соответствуют две ошибки в приеме q_2q_1 , а второй ветви $00 \rightarrow 10$ нуль ошибок. Ошибка по каждой ветви служит метрикой d_H в смысле расстояния Хэмминга, т.е. соответствует числу отличающихся от требуемых принятых бит кода q_2q_1 . Эти метрики указаны на соответствующих переходах в диаграмме (рис.2.7).



Рис. 2.6. Принятая последовательность бит (на входе декодера)

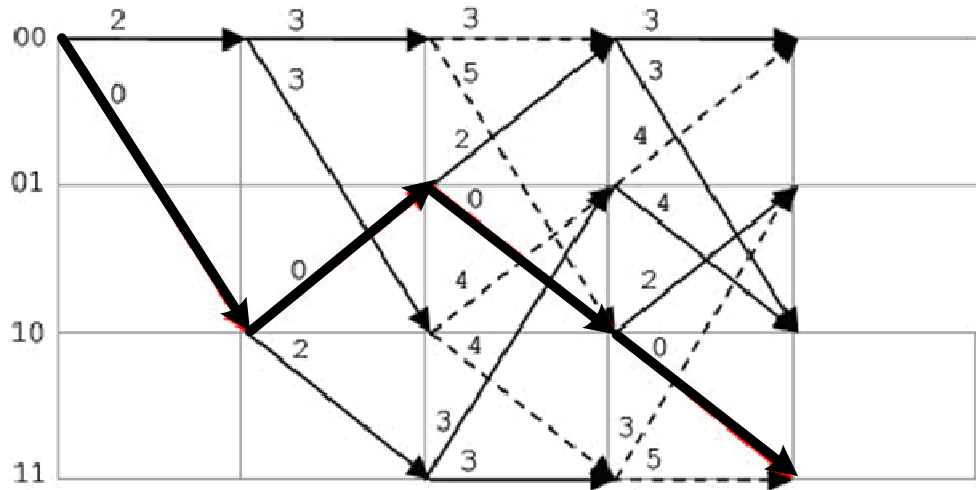


Рис. 2.7. Декодирование на основе алгоритма Витерби.
При приеме ошибки отсутствуют

В момент $z = 2$ сигнал может иметь 4 состояния, которые определяются двумя возможными переходами из 00 и двумя переходами из 10. Сравнение с принятой комбинацией $q_2q_1 = 10$ дает следующие метрики соответствующих ветвей d_H :

$$\begin{aligned} 00 \rightarrow 00 & d_H=1, \\ 00 \rightarrow 10 & d_H=1, \\ 10 \rightarrow 01 & d_H=0, \\ 10 \rightarrow 11 & d_H=2. \end{aligned}$$

Суммарную метрику d_H по каждому из возможных путей определяем как сумму метрик составляющих его ветвей. Значения суммарных метрик показаны на диаграмме рис.2.7 для $z=2$ и последующих моментов времени. Для момента времени $z=3$ следует анализировать уже 8 возможных путей и сравнивать 8 соответствующих им метрик d_H . Алгоритм Витерби выбирает путь с наименьшей суммарной метрикой (штрафом) и начинает отбрасывать по ходу продвижения во времени те пути, которые имеют больший суммарный штраф в каждой точке диаграммы. Поэтому на каждом следующем этапе декодирования в рассматриваемом примере из 8 возможных путей остаются 4. Оптимальным путем является путь с наименьшим штрафом (при отсутствии ошибок в канале связи $d_H=0$). Последовательность бит на выходе декодера, соответствующая этому пути, совпадает с передаваемым информационным сигналом (рис.2.3).

Рассмотрим теперь случай, когда при передаче появились ошибки, например, во втором закодированном бите (рис. 2.8).



Рис. 2.8. Принятая последовательность бит (на входе декодера)

Теперь метрика первой ветви не равна нулю и суммарная метрика оптимального пути $d_H = 1$ (рис. 2.9). Однако, как следует из рис.2.9, оптимальный путь

восстанавливает передачу последовательности информационных бит, т.е. использованный СК исправляет ошибки. Разумеется, алгоритм Витерби не во всех случаях дает верный результат на выходе декодера. Так, при двойной, а тем более тройной последовательной ошибке в передаче кодовых комбинаций путь с наименьшей метрикой d_H содержит ошибку в одном бите на выходе при применении данного кода.

Двойные и тройные ошибки исправляют более мощные коды при использовании полиномов 4, 6 и более высоких порядков. При этом естественно увеличивается объем необходимых вычислений. При большом числе последовательных ошибок (пакетов ошибок) сверточные коды неэффективны. Поэтому перед передачей биты в закодированных кадрах подвергают **перемежению**, чтобы преобразовать при приеме пакеты ошибок в множество единичных ошибок.

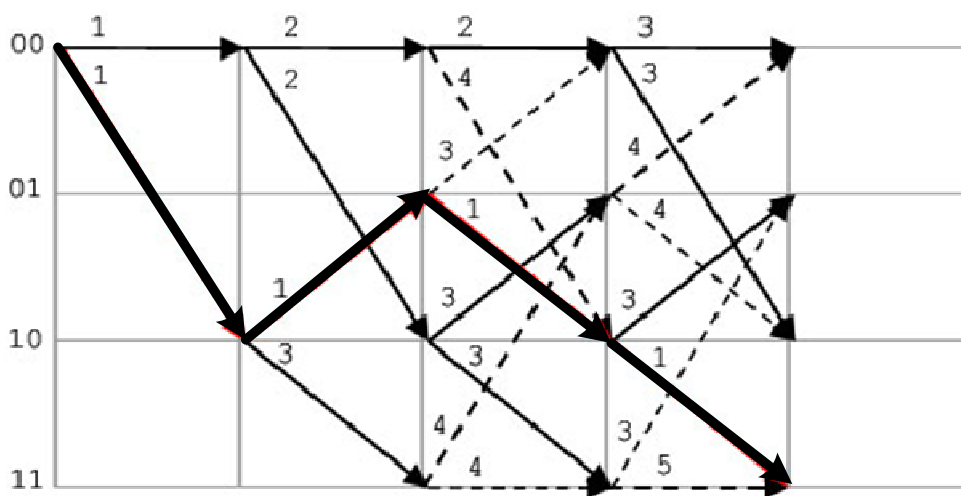


Рис. 2.9. Декодирование на основе алгоритма Витерби.

При приеме есть одна ошибка

Рассмотренный метод декодирования, когда в качестве метрики используют расстояние Хэмминга, называют **декодированием с жестким решением**. Здесь каждому q_i на выходе демодулятора соответствует одно из двух значений: 0 или 1. Лучшие результаты в смысле восстановления исходного сигнала дает **декодирование с мягким решением**. В этом случае все принятые биты кодовых комбинаций на выходе демодулятора подвергают квантованию. Для оптимизации приема сигнала в канале с Гауссовым шумом достаточно использовать 8-уровневые квантователи. Каждому биту кодовой комбинации присваивают при приеме определенный вес. Чем больше этот вес, тем выше вероятность правильного приема. На рис. 2.10 показано распределение весовых коэффициентов для двухпозиционного сигнала⁵. При правильном приеме бита кодовой комбинации вес засчитывают со знаком плюс, при ошибке – со знаком минус.

⁵ Распределение весовых коэффициентов по уровням приема сигналов – прерогатива производителя аппаратуры. При приеме многопозиционных сигналов весовые коэффициенты определяют на основе евклидовых метрик.

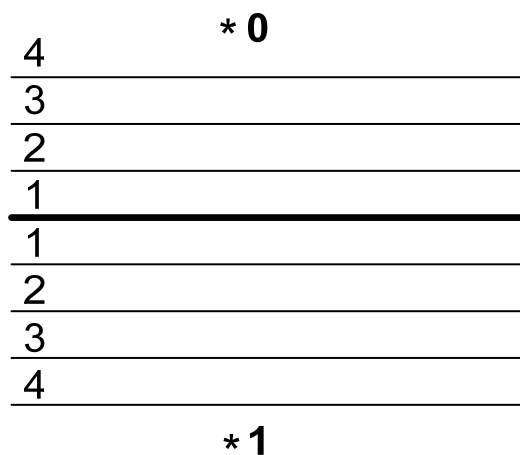


Рис. 2.10. Весовые коэффициенты при приеме двухпозиционного сигнала

На рис. 2.11 и 2.12 рассмотрен вариант декодирования с мягким решением при отсутствии ошибок на входе, а на рис. 2.13 и 2.14 - вариант с наличием ошибок.

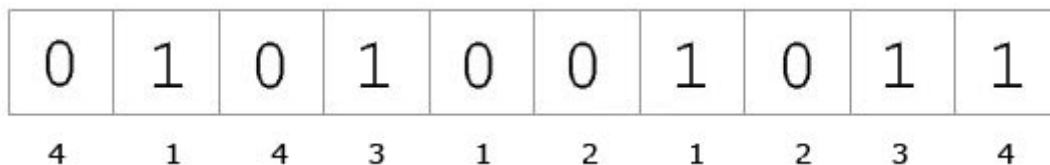


Рис. 2.11. Принятая последовательность бит (на входе декодера)

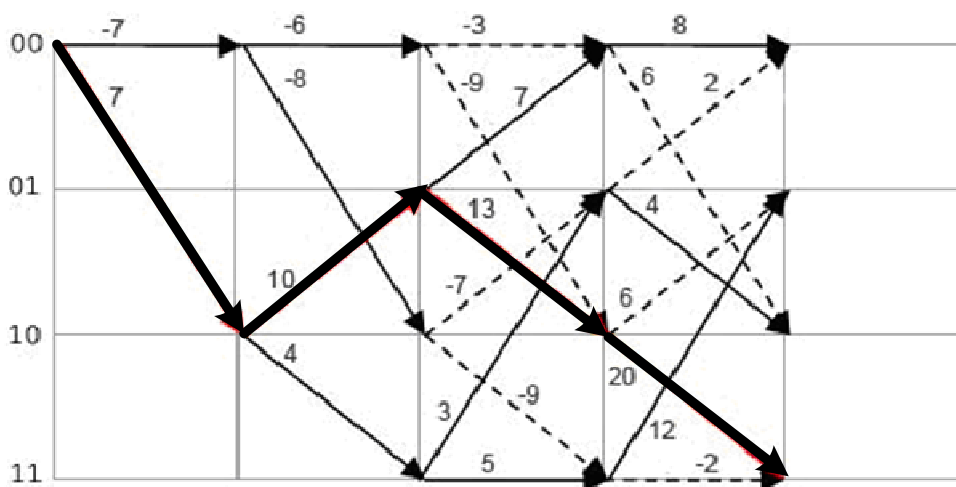


Рис. 2.12. Декодирование с мягким решением.

При приеме ошибки отсутствуют

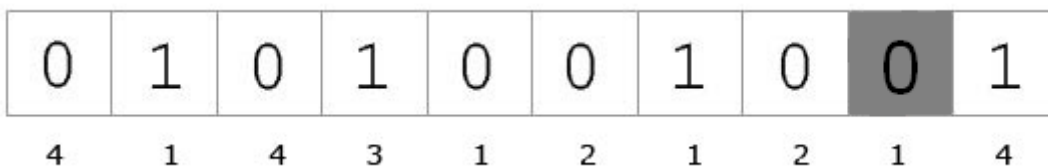


Рис. 2.13. Принятая последовательность бит (на входе декодера)

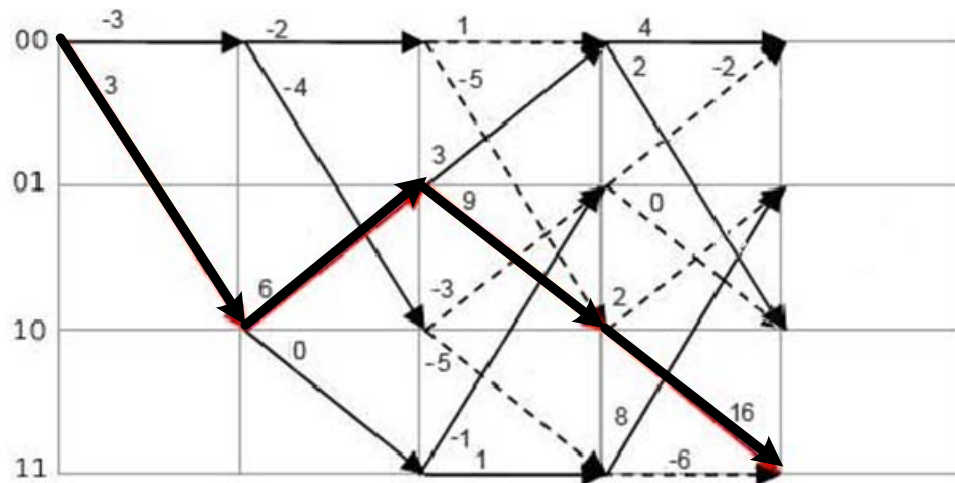


Рис. 2.14. Декодирование с мягким решением.

При приеме есть ошибка

На рис. 2.12 и 2.14 символы q_2q_1 представлены значениями с весовыми коэффициентами, а на ребрах графа приведены соответствующие метрики. Из рис.2.12 следует, что при декодировании с мягким решением в качестве оптимального пути выбирают путь с максимальной суммарной метрикой, что соответствует максимальной накопленной доверительной вероятности. Более того, декодирование с мягким решением, производимое в примерах на рис.2.12, 2.14, при определенных условиях исправляет двойные ошибки в приеме кодовых комбинаций, что соответствует большей помехозащищенности канала связи, чем при использовании декодирования с жестким решением. Это обусловлено тем, что ошибочно принятые биты в среднем имеют меньшие весовые коэффициенты, чем правильно принятые биты.

Рассмотренные примеры позволяют сделать ряд выводов об использовании алгоритма Витерби для декодирования СК. Прежде всего, выбор оптимального пути – это вероятностный процесс и всегда есть хотя бы ничтожно малая вероятность получения ошибочный результат. Вероятность получения правильного результата возрастает с увеличением размерности кода. Однако при этом растут требования к вычислительным способностям (быстродействию) процессора Витерби и объему памяти для хранения промежуточных результатов. На каждом следующем шаге общее число путей возрастает в 2 раза и на z -м шаге при $k = 1$ оно составит 2^{z+1} . Однако в процессе вычислений один из двух путей, сходящихся в один узел, отбрасывают, оставляя путь с наибольшим весом (наименьшим штрафом). Поэтому, начиная с $z = K$, всякий раз отбрасывают половину путей. Следовательно, алгоритм Витерби просчитывает на каждом этапе вычисления метрик не более 2^k путей.

Характеризуя различные сверточные коды, часто говорят об их мощности, понимая под этим способность кодов исправлять множественные, одиночные и многократные ошибки, возникающие в канале связи. Мощность кода зависит от длины кодового ограничения K и от вида образующих полиномов. Можно показать, что вероятность исправления ошибок при декодировании связана со свободным кодовым расстоянием d_f . Величину d_f определяют как кодовое расстояние по Хэммингу между правильным путем с нулевой метрикой (при приеме кодовых комбинаций без ошибок) и ближайшим альтернативным путем (тоже с нулевой метрикой, но с ошибками на входе декодера и на его выходе). Так, для рассмот-

ренного ранее в примере кода (2,1,3) с образующими полиномами q_1 и q_2 (2.5) и (2.6) величина $d_f=5$. Для кода (2,1,5) с образующими полиномами

$$q_1=1+z^3+z^4 \quad (2.12)$$

$$q_2=1+z+z^3+z^4, \quad (2.13)$$

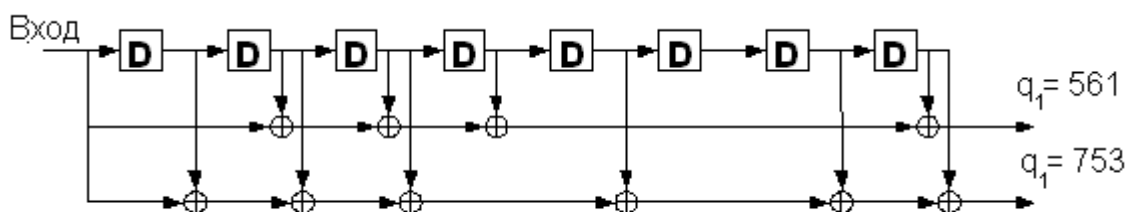
который используют при кодировании речевого сигнала в сетях GSM, $d_f=7$. При декодировании с жестким решением допустимая кратность ошибок символов ограничена $m_q < d_f / 2$. Следовательно, коды с большим кодовым ограничением k являются более мощными.

На рис.2.15 представлены генераторы кодовых комбинаций при кодировании пользовательского и сигнального трафика в стандарте UMTS [17]. В этом стандарте в качестве базовых СК используют коды с $R_{код} = 1/2$ (рис. 2.14 а) и $R_{код} = 1/3$ (рис. 2.14б). Для кодера с $R_{код} = 1/2$ образующие полиномы

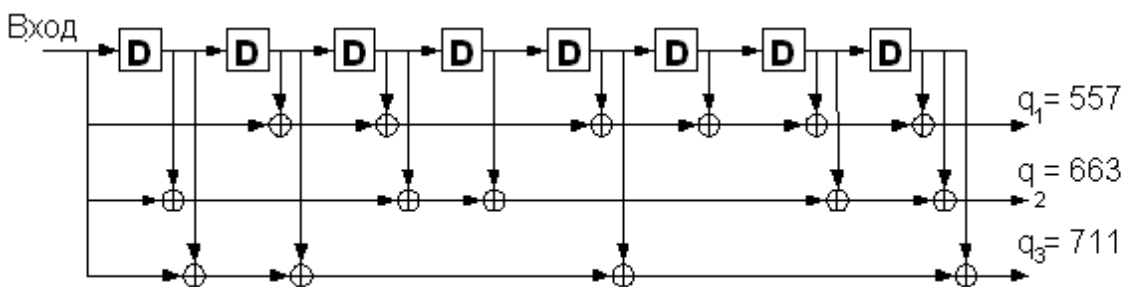
$$q_1=1+z^2+z^3+z^4+z^8 \quad (2.14)$$

$$q_2=1+z+z^2+z^3+z^5+z^7+z^8 \quad (2.15)$$

На рис.2.14 образующие полиномы записаны в цифровом восьмеричном коде, что часто используют в литературе. Следуя справа налево, от последнего бита в сдвигающем регистре (z^8) до входного бита ($z^0=1$), представляем отводы от каждых 3 последовательных ячеек (3 информационных бит) восьмеричным числом.



а) $R_{код} = 1/2$



б) $R_{код} = 1/3$

Рис. 2.15. Структура кодеров СК в стандарте UMTS

Высокая восстанавливающая способность СК позволяет применять на практике СК с перфорацией. Перфорация означает, что из последовательности кодовых комбинаций на выходе кодера часть бит выбрасывают (не передают). При малом числе ошибок в приемнике (каналах с высоким отношением сигнал/помеха) неопределенность, возникающую из-за пропущенных бит, устраняют при вычислении метрик следующих ветвей для двух равновероятных значений

пропущенного бита: $p|_0 = p|_1 = 0,5$. При декодировании с мягким решением весовой коэффициент “выколотых” бит равен 0. В СК с перфорацией повышается скорость передачи. Так, перфорируя (выкалывая) каждый четвертый бит в СК (2,1,3) или (2,1,5), повышают кодую скорость с $R_{кода} = 1/2$ до $R_{кода} = 2/3$.

Степень качества СК зависит от выбранного кода и характеристики канала передачи. Итоговую оценку производят по величине коэффициента ошибок (BER) при заданном отношении сигнал/помеха в канале и типе канала. Поскольку СК хорошо исправляют одиночные ошибки, но чувствительны к пакетам ошибок, то после сверточного кодирования обязательно следует процедура перемежения.

2.4. Перемежение

Перемежение – изменение порядка следования бит внутри определенного блока при его передаче. При приеме блока выполняют операцию **деперемежения**. В результате возможные пакеты ошибок превращаются в некоторое число единичных ошибок, эффективно исправляемых сверточным декодером. Множественные ошибки (пакеты) возникают, прежде всего, из-за быстрых замираний сигнала в точке приема. Пакеты ошибок сложно исправить с помощью блочного кодирования и невозможно – сверточным кодированием. Однако пакетов ошибок можно избежать, если до передачи их переупорядочить в пределах блока, а на стороне приема восстановить исходную последовательность. Разумеется, чем длиннее блок, тем на большее расстояние можно разнести следующие друг за другом биты и тем меньше будет вероятность пакетов ошибок. Однако при этом возрастают задержки при выполнении процедур перемежения и деперемежения. Поэтому при выполнении перемежения всегда ищут компромисс между допустимой задержкой в передаче информации и снижении коэффициента ошибок.

В системах связи используют 4 вида перемежений:

диагональное,
 блоковое,
 межблоковое,
 сверточное [2], [17].

При **диагональном перемежении** блок, состоящий из n бит, разбивают на K сегментов, каждый из которых содержит $m = n/K$ бит. Например, при передаче речи в стандарте GSM 456 бит разбивают на 8 сегментов по 57 бит в каждом (рис.2.16). Далее, как показано на рис.2.16, в кадрах последовательно передают попарно сегменты ($i-1$)-го и i -го блоков, i -го и $(i+1)$ -го блоков и т.д.

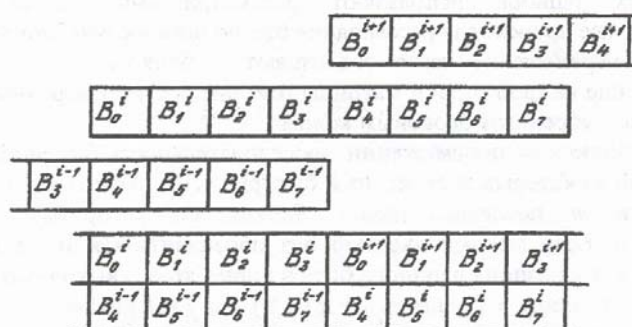


Рис.2.16. Диагональное перемежение

Возможен и иной подход к выполнению диагонального перемежения: используют следующую перекомпоновку сегментов:

$$B_4^{i-1} B_i^i B_5^{i-1} B_2^i B_6^{i-1} B_3^i B_7^{i-1} B_0^i B_4^i B_1^{i+1} B_5^i B_2^{i+1} B_6^i B_3^{i+1} B_7^i \text{ и т.д.}$$

При **блоковом перемежении** последовательность из n бит записывают в виде матрицы, состоящей из K строк и m столбцов.

Например, при $n=456$, $K=8$ и $m=57$ получаем матрицу для $a_0 \dots a_{455}$ бит, приведенную на рис.2.17. Далее считывание производят построчно, что дает возможность сформировать 8 информационных пакетов.

При пропадании даже целого пакета информации (57 бит) использование сверточного кодирования позволяет восстановить исходный информационный сигнал, так как неприятые биты будут разнесены друг от друга на 8 бит.

a_0	a_8	a_{16}	...	a_{448}
a_1	a_9	a_{17}	...	a_{449}
a_2	a_{10}	a_{18}	...	a_{450}
a_3	a_{11}	a_{19}	...	a_{451}
a_4	a_{12}	a_{20}	...	a_{452}
a_5	a_{13}	a_{21}	...	a_{453}
a_6	a_{14}	a_{22}	...	a_{454}
a_7	a_{15}	a_{23}	...	a_{455}

Рис.2.17. Расположение элементов при блоковом перемежении

Следует, однако, отметить, что если ошибки при приеме бит имеют периодичность, совпадающую с длиной столбца на рис.2.17, то появляются пакеты ошибок, которые невозможно исправить сверточными кодами. Для устранения таких периодических ошибок используют межблоковое перемежение, где происходит рассеивание бит по псевдослучайному закону.

При **межблоковом перемежении** последовательность бит разбивают на блоки, каждый из которых состоит из u бит (рис.2.18), и биты каждого блока рассеивают в m последовательных блоках по алгоритму, где нет периодичности. Если каждый входной бит обозначить как $x(i,k)$, где i - номер блока, а k - позиция входного бита в блоке, то его выходную позицию определяют по правилу $y(i+j, j+mt) = x$, где $j = k \bmod m$, $t = k \bmod N$, и в каждом блоке $n = mN$ бит.

В примере, проиллюстрированном рис.2.17, $m = 3$, $N = 2$, $n = 6$. Как следует из рис.2.18, периодичность в распределении бит устранена, однако перемежение бит каждого блока растягивается на 3 блока (в общем случае на m блоков), что вводит дополнительную задержку при передаче.

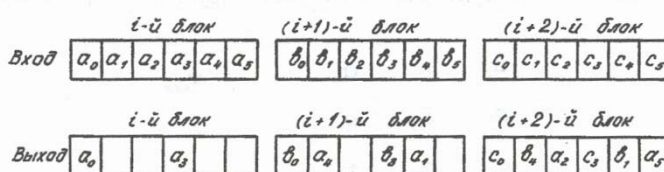


Рис.2.18. Межблоковое перемежение

Сверточное перемежение напоминает блочное. Так же, как и при блочном перемежении, формируют матрицу, но процесс считывания бит ведут не по строкам, а по диагоналям матрицы (рис.2.19). При сверточном перемежении ослаблено влияние периодических ошибок и требуется меньший объем памяти, чем при блочном.

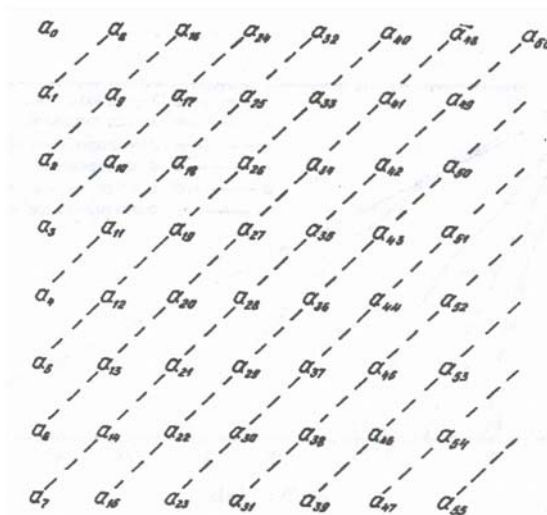


Рис.2.19. Траектории считывания элементов при сверточном перемежении

В стандартах радиодоступа широко применяют комбинации различных методов перемежения, добиваясь псевдослучайного расположения бит кодированных сигналов в передаваемых по радиоканалам блоках.

Контрольные вопросы к главе 2

1. С какой целью используют избыточное кодирование в сетях радиодоступа.
2. Как осуществляют обнаружение ошибок при передаче кадров (блоков) данных.
3. Что определяет "расстояние по Хэммингу". Как оно связано с обнаружением и исправлением ошибок.
4. Опишите алгоритм блочного кодирования.
5. Что определяет синдром при блочном кодировании и как его найти.
6. Как определяют контрольную сумму (CRC) при передаче кадров (блоков) данных.
7. Опишите алгоритм сверточного кодирования. Почему он позволяет обнаруживать и исправлять ошибки.
8. Какой вероятностный алгоритм реализует алгоритм Витерби.
9. По какому критерию определяют оптимальный путь в алгоритме Витерби с жестким решением.
10. По какому критерию определяют оптимальный путь в алгоритме Витерби с мягким решением.
11. Сколько состояний будет в решетчатой диаграмме сигнала, если при сверточном кодировании используют образующие полиномы n -го порядка.
12. Ошибки какой кратности могут быть исправлены, если при сверточном кодировании используют образующие полиномы n -го порядка.
13. Зачем после кодирования информационных кадров выполняют операцию перемежения.
14. В чем состоит алгоритм блочного перемежения.
15. Опишите алгоритм межблокового перемежения.

3. Адаптивные эквалайзеры в приемниках радиоустройств стандарта GSM

Ошибки при приеме цифровых сигналов обусловлены как влиянием помех, так и межсимвольной интерференцией – перекрытием символов из-за ограниченной полосы канала связи и многолучевого приема. При ограничении полосы сигнала вместо прямоугольных импульсов или половин синусоид появляются символы теоретически бесконечной длительности, перекрывающие соседние символы (рис. 3.1).

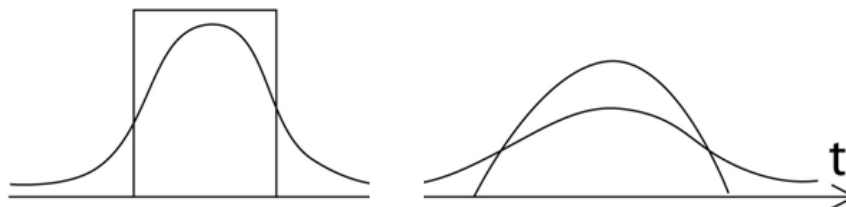


Рис.3.1. Растяжение символов при ограничении полосы канала связи

Устройства в приемнике, минимизирующие эффект межсимвольной интерференции, называются эквалайзерами. При известных характеристиках канала связи эквалайзеры реализуют на основе трансверсальных цифровых фильтров, но присутствие помех усложняет оптимальные алгоритмы приема [1], [4], [14], [15].

При многолучевом распространении сигнала (рис. 3.2) на входе приемника происходит сложение множества лучей со случайными взаимными временными сдвигами, амплитудами и фазами приходящих сигналов. Схематически это можно проиллюстрировать рис.3.3.

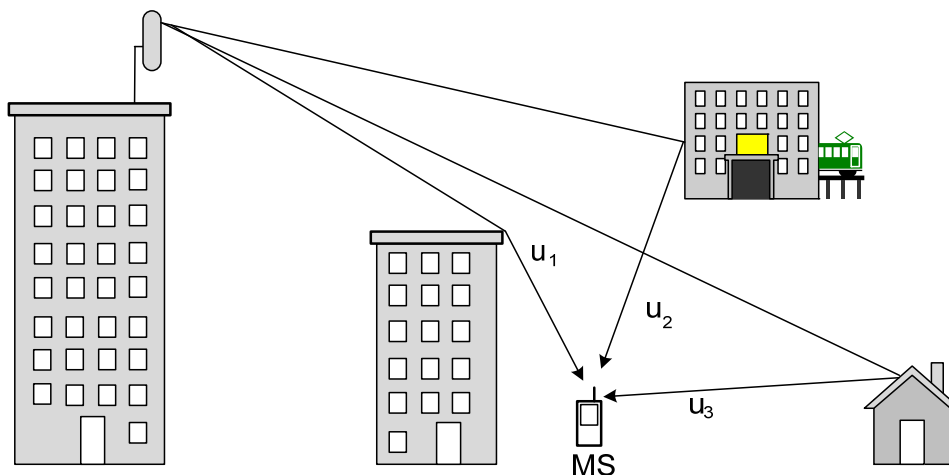
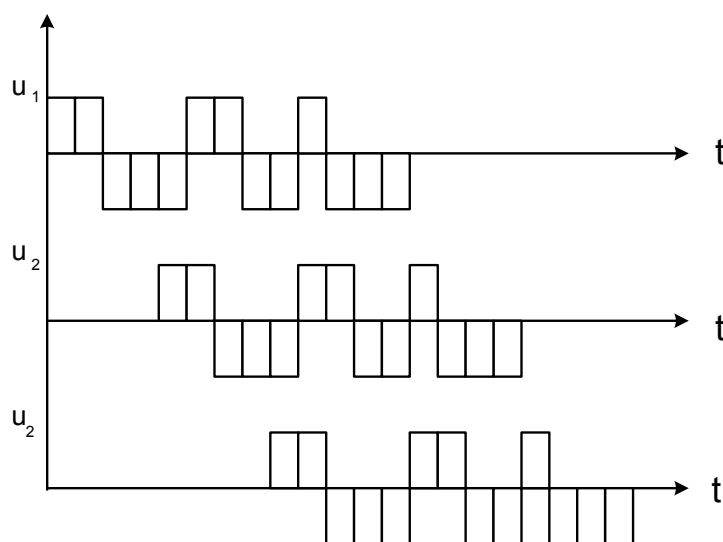


Рис.3.2. Многолучевое распространение сигналов



*Рис.3.3. Межсимвольная интерференция
из-за многолучевого распространения сигналов*

С повышением скорости передачи разности по времени прихода лучей может достигать десятков длительностей символов. В стандартах радиодоступа 21 века используют 2 варианта решения этой проблемы. Первый вариант состоит в выделении и отдельной обработке приходящих на приемник лучей, что возможно при когерентном приеме в системах с кодовым разделением каналов: в сетях стандартов UMTS, CDMA2000. Второй вариант – применение технологии OFDM, когда сигнал по радиоканалу передают на множестве поднесущих частот “длинными” символами, разделенными защитными промежутками [1], [18]. Так организованы физические каналы в стандартах Wi-Fi, WiMAX, LTE.

Как правило, в спецификациях стандартов радиодоступа нет указаний на использование эквалайзеров в приемниках абонентских терминалов и базовых станций. Эквалайзеры разрабатывают и используют сами производители аппаратуры для улучшения качества приема радиосигналов. Исключения составляют стандарты GSM и TETRA, где применение эквалайзеров обязательно. Рассмотрим принципы работы и структуру эквалайзеров в приемной аппаратуре сетей стандарта GSM [18].

Информацию по радиоканалам сетей GSM передают в виде пакетов (burst). Структура пакета данных стандарта GSM (normal burst) приведена на рис. 3.4.

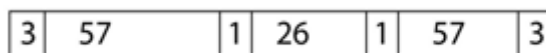


Рис.3.4. Пакет данных стандарта GSM (normal burst)

В начале и в конце пакета размещены 3 символа (“хвосты”), необходимые для выхода передатчика на требуемую мощность и его выключения. Далее следуют 2 блока кодированных данных по 57 символов в каждом. Ближе к центру находятся 2 флажка (по 1 символу), определяющих, что передано в полях данных: трафик или сигнализация. В центре пакета – 26 символов, составляющих зондирующий сигнал, называемый обучающей последовательностью (training sequence). Математически зондирующий сигнал представляет собой m-последовательность, имеющую один максимум автокорреляционной функции. В [19] приведены структуры обучающих последовательностей для различных видов

модуляции, используемой в сетях GSM. Для двухпозиционного сигнала гауссовской ЧММС (частотной модуляции с минимальным сдвигом) специфицированы 2 конфигурации обучающих последовательностей по 8 вариантов в каждой.

Конфигурация 1:

0	(0,0,1,0,0,1,0,1,1,1,0,0,0,0,1,0,0,0,1,0,0,1,0,0,1,0,1,1,1)
1	(0,0,1,0,1,1,0,1,1,1,0,1,1,1,1,0,0,0,1,0,1,1,0,1,1,1)
2	(0,1,0,0,0,0,1,1,1,0,1,1,1,0,1,0,0,1,0,0,0,0,1,1,1,0)
3	(0,1,0,0,0,1,1,1,1,0,1,1,0,1,0,1,0,0,0,1,0,0,0,1,1,1,0)
4	(0,0,0,1,1,0,1,0,1,1,1,0,0,1,0,0,0,0,0,1,1,0,1,0,1,1)
5	(0,1,0,0,1,1,1,0,1,0,1,1,0,0,0,0,0,1,0,0,1,1,1,0,1,0)
6	(1,0,1,0,0,1,1,1,1,0,1,1,0,0,0,1,0,1,0,0,1,1,1,1,1)
7	(1,1,1,0,1,1,1,1,0,0,0,1,0,0,1,0,1,1,1,0,1,1,1,1,0,0)

Конфигурация 2:

0	(0,1,1,0,0,0,1,0,0,0,1,0,0,1,0,0,1,1,1,1,0,1,0,1,1,1)
1	(0,1,0,1,1,1,1,0,1,0,0,1,1,0,1,1,1,0,1,1,1,0,0,0,0,1)
2	(0,1,0,0,0,0,0,1,0,1,1,0,0,0,1,1,1,0,1,1,1,0,1,1,0,0)
3	(0,0,1,0,1,1,0,1,1,1,0,1,1,1,0,0,1,1,1,1,0,1,0,0,0,0)
4	(0,1,1,1,0,1,0,0,1,1,1,1,0,1,0,0,1,1,1,0,1,1,1,1,0)
5	(0,1,0,0,0,0,0,1,0,0,1,1,0,1,0,1,0,0,1,1,1,1,0,0,1,1)
6	(0,0,0,1,0,0,0,0,1,1,0,1,0,0,0,0,1,1,0,1,1,1,0,1,0,1)
7	(0,1,0,0,0,1,0,1,1,1,0,0,1,1,1,1,1,1,0,0,1,0,1,0,0,1)

Фактически зондирующий сигнал состоит из 16 центральных бит 26-битовой последовательности; 5 бит слева и 5 бит справа введены для устранения межсимвольной интерференции со случайными битами информационных полей и флажков. Далее рассмотрим принципы реализации адаптивного эквалайзера GSM при гауссовской ЧММС. Этот вид модуляции является основным при передаче телефонного трафика.

В сетях GSM допустима задержка при многолучевом распространении сигналов до 15 мкс, что при сквозной скорости передачи в радиоканале $B_{скв} = 271$ кбит/с соответствует задержке до 4 бит (длительность бита $T_B = 1/B_{скв} = 3,69$ мкс). Из-за такой межсимвольной интерференции разработчики стандарта GSM сузили полосу радиоканала до 200 МГц, что приводит к перекрытию каждого информационного символа двумя соседними (рис. 3.5).

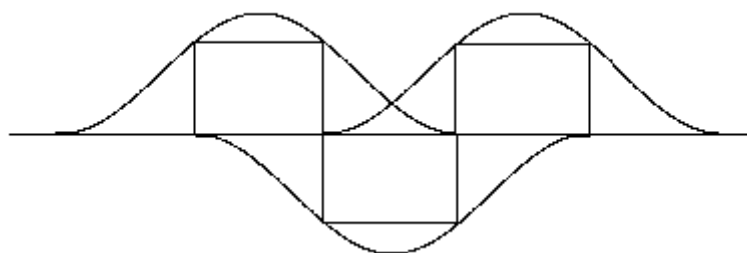


Рис.3.5. Перекрытие информационных символов из-за сужения полосы канала

Напомним, что сигнал гауссовской ЧММС комплексный, состоящий из синфазной и квадратурной составляющих:

$$u_{чммс} = U \left(\cos \frac{\pi t}{2T_B} \sin \omega_\phi t \pm \sin \frac{\pi t}{2T_B} \cos \omega_\phi t \right) \quad (3.1)$$

Процесс межсимвольной интерференции может быть описан как свертка последовательности информационных символов (половина синусоиды при ЧММС) и импульсной характеристики фильтра, ограничивающего полосу передачи (рис. 3.6):

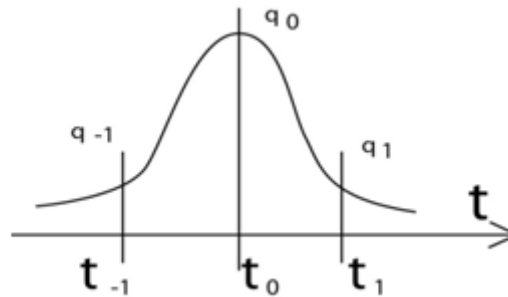


Рис.3.6. Импульсная характеристика фильтра ограничения полосы

В момент времени $t=0$, для текущего символа a_0 и каждого из трех предыдущих имеем:

$$\begin{aligned}
 \tilde{a}_0 &= \tilde{u}_1 q_{-1} + \tilde{u}_0 q_0 + \tilde{u}_{-1} q_1 \\
 \tilde{a}_{-1} &= \tilde{u}_0 q_{-1} + \tilde{u}_{-1} q_0 + \tilde{u}_{-2} q_1 \\
 \tilde{a}_{-2} &= \tilde{u}_{-1} q_{-1} + \tilde{u}_{-2} q_0 + \tilde{u}_{-3} q_1 \\
 \tilde{a}_{-3} &= \tilde{u}_{-2} q_{-1} + \tilde{u}_{-3} q_0 + \tilde{u}_{-4} q_1
 \end{aligned}
 \tag{3.2}$$

Поскольку ограничение полосы при ЧММС при передаче сигналов осуществляют стандартным фильтром, то отсчеты его импульсной характеристики q_{-1}, q_0 и q_1 известны, так что полученные символы $a_0, a_{-1}, a_{-2}, a_{-3}$ также известны. При этом все a_i и u_i представляют собой комплексные величины.

Процесс многолучевого распространения сигнала на радиоинтерфейсе можно рассматривать как импульсную реакцию эквивалентного фильтра (рис 3.7), где сплошной кривой обозначена вещественная часть передаточной функции, а пунктиром — мнимая часть.

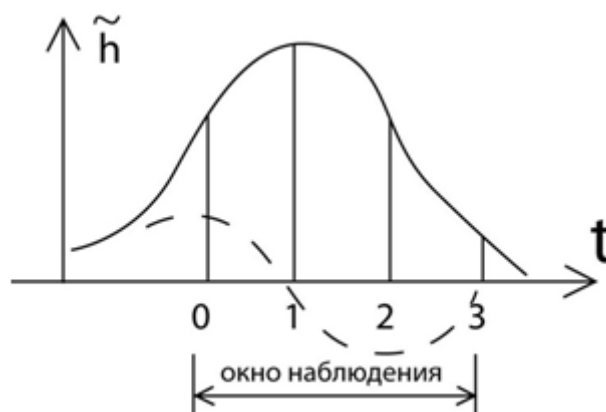


Рис.3.7. Характеристика эквивалентного фильтра радиоканала

В результате сигнал на входе GSM-приемника:

$$\tilde{r}_0 = \tilde{a}_0 \tilde{h}_0^W + \tilde{a}_{-1} \tilde{h}_1^W + \tilde{a}_{-2} \tilde{h}_2^W + \tilde{a}_{-3} \tilde{h}_3^W
 \tag{3.3}$$

Учитывая (3.2), получаем

$$\begin{aligned}
\tilde{r}_0 &= (\tilde{u}_1 q_{-1} + \tilde{u}_0 q_0 + \tilde{u}_{-1} q_1) \tilde{h}_0^W + \\
&(\tilde{u}_0 q_{-1} + \tilde{u}_{-1} q_0 + \tilde{u}_{-2} q_1) \tilde{h}_1^W + \\
&(\tilde{u}_{-1} q_{-1} + \tilde{u}_{-2} q_0 + \tilde{u}_{-3} q_1) \tilde{h}_2^W + \\
&(\tilde{u}_{-2} q_{-1} + \tilde{u}_{-3} q_0 + \tilde{u}_{-4} q_1) \tilde{h}_3^W
\end{aligned} \tag{3.4}$$

Под \tilde{h}_i^W понимаем отсчеты комплексной передаточной функции \tilde{h}_i , в окне наблюдения, соответствующем в GSM длительности 4 бит (символов). Сложность анализа заключается в том, что отсчеты импульсной характеристики \tilde{h}_i^W заранее неизвестны, и в каждый момент случайны. Их можно определить только экспериментально. Для этого и используют зондирующий сигнал $C(t)$ обучающей последовательности в пакете данных, передаваемых по радиоканалу.

Определение импульсной характеристики сквозного канала (сквозной канал представляют в виде каскадного соединения фильтра ограничения полосы передачи и эквивалентного фильтра радиоканала) проводят следующим образом.

В приемнике зондирующий сигнал (16 бит+5+5 бит “хвостов”) пропускают через согласованный фильтр. Принятый зондирующий сигнал представляет собой свертку:

$$\tilde{C}_{RX} = \tilde{C}(t) ** \tilde{h}_3(t) \tag{3.5}$$

где $\tilde{C}(t) = C_I(t) + jC_Q(t)$ – исходный комплексный зондирующий сигнал, а $\tilde{h}_3(t)$ – импульсная характеристика сквозного канала.

Операция согласованной фильтрации состоит в свертке сигнала $\tilde{C}(t)$ и комплексно-сопряженного сигнала $\tilde{C}^*(t)$:

$$\begin{aligned}
\rho(t) &= \tilde{C}(t) ** \tilde{C}^*(t) = [C_I(t) + jC_Q(t)] ** [C_I - jC_Q(t)] = \\
&= C_I(t) ** C_I(t) + C_Q(t) ** C_Q(t)
\end{aligned} \tag{3.6}$$

В результате, $\rho(t)$ представляет собой заранее известную реальную функцию неопределенности, ширина всплеска которой составляет $2T_B$.

В приемнике GSM согласованную фильтрацию осуществляют сигналом (3.5):

$$\tilde{C}_{ex} ** \tilde{C}^*(t) = \tilde{C}(t) ** \tilde{h}_3(t) ** \tilde{C}^*(t) = \rho(t) ** \tilde{h}_3(t) \tag{3.7}$$

В результате на выходе коррелятора зондирующего сигнала получаем свертку известной функции неопределенности $\rho(t)$ и импульсной характеристики сквозного канала. Ограничивая размеры временной зависимости размерами временного окна, получаем временную зависимость $\rho(t) ** \tilde{h}_3^W(t)$ для анализа принятого сигнала и синтеза переданного.

Структура адаптивного эквалайзера, реализованного в GSM-приемниках, приведена на рис. 3.8.



Рис.3.8. Структура адаптивного эквалайзера GSM-приемника

Из принятого пакета (burst) в демультимплексоре выделяют два информационных блока (по 57+1 бит) и обучающую последовательность. Ее направляют в блок оценки канала, где подвергают согласованной фильтрации. В результате на выходе блока получают отсчеты импульсной характеристики сквозного канала $\rho(t) ** \tilde{h}_3^w(t)$.

Каждый из информационных блоков по 58 бит свертывают с функцией $\rho(t)$, в результате чего получают сигнал

$$\tilde{X}(t) = \tilde{r}(t) ** \rho(t) \quad (3.8)$$

Следующий этап работы эквалайзера состоит в синтезе всех возможных вариантов принятого сигнала, для чего используют генератор псевдослучайных последовательностей. В каждый тактовый момент времени на выходе генератора ПСП появляются 2^6 вариантов последовательностей длиной в 6 бит, что соответствует числу бит, учитываемых в выражении (3.4).

Эти 2^6 вариантов в блоке генератора псевдослучайных последовательностей подвергают свертке с рассчитанной функцией $\rho(t) ** \tilde{h}_3^w(t)$, получая 2^6 функций $\tilde{X}'(t)$. В блоке вычисления инкрементов метрик вычисляют 2^6 разностей $|\tilde{X}(t) - \tilde{X}'(t)|$. Наиболее правдоподобному результату соответствует минимум метрики $d(t) = |\tilde{X}(t) - \tilde{X}'(t)|$. Из 64 найденных решений оставляют 32 варианта с минимальными метриками.

Этот процесс проводят от 1-го бита информационного блока для последнего, используя для нахождения наиболее вероятного варианта переданного сигнала алгоритм Витерби, рассмотренный в гл. 2 пособия..

Алгоритм включает в себя 32 вероятных состояния сигнала. При определении очередного бита из каждого состояния имеется два возможных перехода. Оптимальным считают путь с минимальной суммарной метрикой, на основе которого и восстанавливают (синтезируют) переданный сигнал.

Контрольные вопросы к главе 3

1. Какие функции выполняют эквалайзеры в приемных устройствах аппаратуры радиодоступа.
2. Почему в приемниках аппаратуры сетей GSM необходимо использовать эквалайзеры.
3. Что такое обучающая последовательность и как ее используют в адаптивных эквалайзерах.
4. Как определяют импульсную характеристику сквозного канала.
5. На основе каких алгоритмов получают (синтезируют) принятый приемником GSM сигнал.

СПИСОК ЛИТЕРАТУРЫ

1. Бабков В.Ю., Цикин И.А. Сотовые системы мобильной радиосвязи.- СПб.: БХВ-Петербург, 2013.- 432 с.
2. Бабков В.Ю., Вознюк М.А., Петраков В.А., Рыжков А.Е., Сиверс М.А. Передача информации в системах подвижной связи. - СПб: ГУТ, 1999.- 152с.
3. Рихтер С.Г. Кодирование и передача речи в цифровых системах подвижной связи.- М.: Горячая линия-телеком, 2009.- 304с.
4. Варгаузин В.А., Цикин И.А. Методы повышения энергетической и спектральной эффективности цифровой радиосвязи. –Спб.: БХВ-Петербург, 2013. – 352с.
5. 3GPP TS 26.090. V11.0.0 Adaptive Multi-Rate (AMR) speech codec; Transcoding functions
6. 3GPP TS 06.10. V8.2.0 Full rate speech; Transcoding
7. ETSI EN 300 969 V8.0.1 Half rate speech transcoding
8. 3GPP TS 26.190 V11.0.0 Adaptive Multi-Rate - Wideband (AMR-WB) speech codec; Transcoding functions
9. 3GPP TS 26.290 V11.0.0 Extended Adaptive Multi-Rate - Wideband (AMR-WB+) speech codec; Transcoding functions
10. Волков А.Н., Рыжков А.Е., Сиверс М.А. UMTS. Стандарт сотовой связи третьего поколения. – СПб, Линк, 2008.- 224с.
11. 3GPP TS 26.071. V11.0.0 Mandatory speech CODEC speech processing functions; AMR speech CODEC; General description
12. 3GPP TS 26.094. V11.0.0 Mandatory speech codec speech processing functions; Adaptive Multi-Rate (AMR) speech codec; Voice Activity Detector (VAD)
13. 3GPP TS 26.092. V11.0.0 Mandatory speech codec speech processing functions; Adaptive Multi-Rate (AMR) speech codec; Comfort noise aspects
14. Скляр Б. Цифровая связь. Теоретические основы и практическое применение. М.: Вильямс. 2003.- 1105с.
15. Прокис Дж. Цифровая связь. – М.: Радио и связь, 2000.- 800с.
16. Р. Морелос-Сарагоса. Искусство помехоустойчивого кодирования. Методы, алгоритмы, применение. Пер. с англ., М.: Техносфера, 2005.- 320с.
17. 3GPP TS 25.212. V11.0.0 Multiplexing and Channel Coding (FDD)
18. Mobile Radio Communications /Raymond Steele (Ed).- L.:Pentech Press Publishers, 1992.- 779p.
19. 3GPP TS 45.002. V12.0.0 Multiplexing and multiple access on the radio path

СОДЕРЖАНИЕ

Введение	3
1. Технологии передачи телефонии в сетях подвижной связи	4
1.1. Основы цифровой передачи телефонии.....	4
1.2. Снижение скорости передачи речи. Дифференциальная ИКМ.....	7
1.3. Дельта-модуляция.....	9
1.4. Эффективное кодирование речи. Принципы сжатия речевых сигналов	10
1.5. Обработка речи в вокодерах	13
1.6. Вокодер AMR	17
1.7. Развитие техники передачи речи в сотовых сетях связи	30
1.8. Прерывистая передача речи	35
2. Кодирование информационных сигналов.....	40
2.1. Необходимость кодирования информационных сигналов.....	40
2.2. Блочное кодирование.....	40
2.3. Сверточное кодирование.....	45
2.4. Перемежение.....	53
3. Адаптивные эквалайзеры в приемниках радиоустройств	
стандарта GSM	54
Список литературы.....	63