

4.11. Кратковременная функция среднего разности значений (КФСР) сигнала $x(n)$ определяется выражением [см. (4.38)]

$$\gamma_n(k) = \frac{1}{N} \sum_{m=0}^{N-1} |x(n+m) - x(n+m-k)|.$$

а) Используя неравенство [16]

$$\frac{1}{N} \sum_{m=0}^{N-2} |x(m)| \leq \left[\frac{1}{N} \sum_{m=0}^{N-1} |x(m)|^2 \right]^{1/2}.$$

показать, что $\gamma_n(k) \leq [2(R_n(0) - R_n(k))]^{1/2}$. Этот результат приводит к равенству (4.39).

б) Изобразить $\gamma_n(k)$ и величину $[2(R_n(0) - R_n(k))]^{1/2}$ при $0 \leq k \leq 200$ для сигнала $x(n) = \cos(\omega_0 n)$ с $N=200$, $\omega_0 = 200\pi/(10\,000)$.

4.12. Рассмотрим входной сигнал $x(n) = A \cos(\omega_0 n)$ трехуровневого центрального ограничителя вида

$$y(n) = \begin{cases} 1, & x(n) > C_L; \\ 0, & |x(n)| \leq C_L; \\ -1, & x(n) < -C_L. \end{cases}$$

а) Изобразить $y(n)$ как функцию n для $C_L = 0,5A$, $0,75A$ и A .

б) Изобразить автокорреляционные функции для $y(n)$ и значений C_L из п. а).

в) Обсудить влияние взаимного расположения C_L и A . Пусть A изменяется во времени, удовлетворяя неравенству $0 < A(n) \leq A_{max}$. Рассмотреть ситуацию, которая может возникнуть при C_L , близком к A_{max} .

5

Цифровое представление речевых сигналов

5.0. Введение

«Если бы я смог заставить поток электричества изменяться по интенсивности точно в соответствии с изменением плотности воздуха во время распространения в нем звука, я бы смог передавать по телеграфу любые звуки, даже звуки речи» — А. Г. Белл [1]. Эта простая идея, имеющая столь важное значение для истории связи, кажется сегодня очевидной. Принцип, изложенный в открытии Белла, положен в основу множества устройств и систем, предназначенных для записи, передачи или обработки речевых сигналов и в которых речевой сигнал отражает колебания плотности звуковых (речевых) волн. Это от-

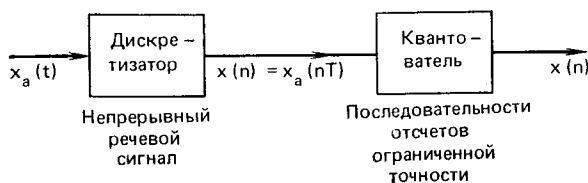


Рис. 5.1. Общая схема цифрового представления

носится и к цифровым системам, в которых речевой сигнал представлен последовательностью своих мгновенных значений.

Общая схема цифрового представления речевого сигнала изображена на рис. 5.1. Из рисунка следует, что речевое колебание как непрерывная функция времени подвергается дискретизации, чаще всего периодической, в результате которой образуется последовательность отсчетов $x_a(nT)$. Эти отсчеты могут в общем случае принимать непрерывное множество значений. Поэтому для получения цифрового, т. е. дискретного по амплитуде и по времени, представления необходимо проквантовать каждый отсчет до конечного множества значений.

Мы увидим далее, что рис. 5.1 достаточно полно отражает процесс формирования цифрового представления речевого сигнала. Может быть не во всех случаях можно разделить эту процедуру на два отдельных этапа, но основные операции — дискретизация и квантование — свойственны всем методам, приведенным в данной главе.

В начале главы изложены вопросы дискретизации применительно к речевым сигналам. Далее излагаются методы квантования отсчетов речевого колебания.

5.1. Дискретизация речевых сигналов

Теорема дискретизации уже обсуждалась в гл. 2. Последовательность отсчетов сигнала, как показано в гл. 2, единственным образом описывает аналоговый сигнал, если он ограничен по полосе частот и частота дискретизации по крайней мере вдвое больше наивысшей частоты спектра сигнала. Поскольку нас интересует цифровое представление речевых сигналов, изучим спектральные свойства речи. В соответствии с изложенным в гл. 3 описанием гласных и фрикативных звуков речевой сигнал не ограничен по полосе частот, хотя его спектр быстро спадает в области высоких частот. На рис. 5.2 изображены спектры типичных звуков речи.

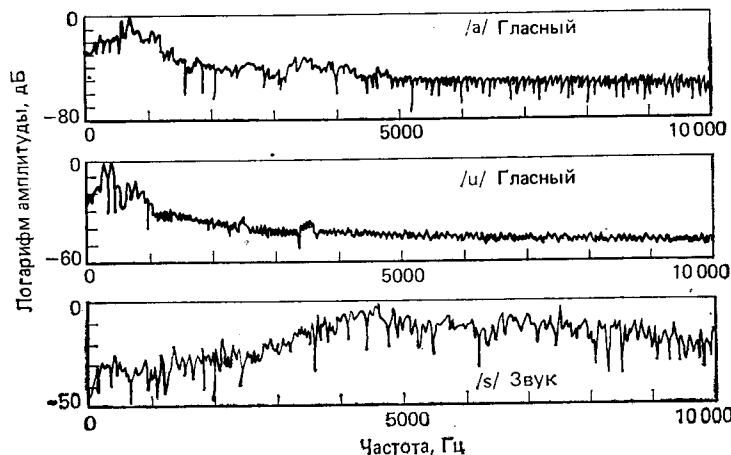


Рис. 5.2. Спектры вокализованных звуков /a/ и /u/ и невокализованного /s/ при частоте дискретизации 20 кГц

Видно, что для вокализованных звуков наивысшая частота, ниже которой максимумы спектра меньше уровня 40 дБ, составляет около 4 кГц. С другой стороны, для невокализованных звуков

спектр не затухает даже на частотах выше 8 кГц. Таким образом, для точного воспроизведения всех звуков речи требуется частота дискретизации около 20 кГц. В большинстве приложений такая частота дискретизации, однако, не требуется. Например, если дискретизация предшествует оцениванию трех первых формантных частот вокализованной речи, то достаточно располагать частью спектра до частоты около 3,5 кГц. Таким образом, если перед дискретизацией речевой сигнал пропускается через фильтр нижних частот так, что частота Найквиста равна 4 кГц, то частота дискретизации должна составлять 8 кГц. Другой пример. Рассмотрим речевой сигнал, который требуется передать по телефонному каналу. На рис. 5.3 приведена типичная частотная характеристика телефонного канала. Очевидно, что телефонный канал ограничивает полосу частот сигнала и частота Найквиста для «телефонной речи» составляет примерно 4 кГц.

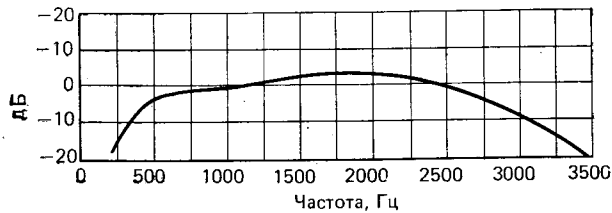


Рис. 5.3. Типичная частотная характеристика тракта телефонной связи (по BTL, Transmission Systems for communication, стр. 73)

Важная особенность, которую часто не замечают при обсуждении дискретизации, состоит в том, что даже если сигнал имеет ограниченный по частоте спектр, он может быть искажен широкополосным случайным шумом перед аналого-цифровым преобразованием. В таких случаях смесь сигнала и шума должна быть пропущена через фильтр с частотой среза, близкой к частоте Найквиста, что позволит избежать эффекта наложения частот при цифровом представлении.

5.2. Обзор статистических моделей речевых сигналов

При рассмотрении цифровых методов представления часто достаточно предполагать, что речевой сигнал является эргодическим случайным процессом. Хотя это является большим упрощением, далее будет показано, что статистическая точка зрения приводит к полезным результатам, тем самым подтверждая целесообразность подобной модели.

Если предположить, что сигнал $x_a(t)$ представляет собой непрерывный случайный процесс, то периодическая последовательность отсчетов этого сигнала может рассматриваться как случайный процесс с дискретным временем. В ряде случаев при анализе систем связи адекватными характеристиками аналогового сигнала являются одномерная функция плотности вероятности и автокорре-

ляционная функция, определенная выражением

$$\varphi_a(\tau) = E[x_a(t)x_a(t+\tau)], \quad (5.1)$$

где $E[\cdot]$ означает усреднение по ансамблю величины, стоящей в квадратных скобках. Непрерывная спектральная плотность мощности представляет собой преобразование Фурье от $\varphi_a(\tau)$:

$$\Phi_a(\Omega) = \int_{-\infty}^{\infty} \varphi_a(\tau) e^{-i\Omega\tau} d\tau. \quad (5.2)$$

Сигнал с дискретным временем, полученный из непрерывного сигнала, имеет автокорреляционную функцию

$$\varphi(m) = E[x(n)x(n+m)] = E[x_a(nT)x_a(nT+mT)] = \varphi_a(mT). \quad (5.3)$$

Это просто дискретизированная функция $\varphi_a(\tau)$, поэтому спектральная плотность мощности равна

$$\begin{aligned} \Phi(e^{i\Omega T}) &= \sum_{m=-\infty}^{\infty} \varphi(m) e^{-i\Omega T m} = \\ &= \frac{1}{T} \sum_{k=-\infty}^{\infty} \Phi_a\left(\Omega + \frac{2\pi}{T}k\right). \end{aligned} \quad (5.4)$$

Из (5.4) следует, что спектральная плотность дискретизированного сигнала представляет собой периодическую последовательность, каждый член которой повторяет спектр аналогового сигнала.

Функция плотности вероятности величины $x(n)$ такая же, как и величины $x_a(t)$, так как $x(n) = x_a(nT)$. Это означает, в свою очередь, что и среднее и дисперсия непрерывного сигнала и сигнала с дискретным временем одинаковы.

Для использования статистических понятий при описании речевых сигналов необходимо оценить функцию плотности вероятности и корреляционную функцию (или спектральную плотность мощности) речевого колебания. Функция плотности вероятности оценивается путем определения гистограммы по большому числу отсчетов, т. е. в течение большого отрезка времени. Давенпорт [2] провел обширные исследования такого рода, а позже Паез и Глисон [3], используя сходные измерения, показали, что хорошей аппроксимацией для экспериментальной функции плотности вероятности может служить гамма-распределение

$$p(x) = \left(\frac{\sqrt{3}}{8\pi\sigma_x|x|}\right)^{1/2} e^{\frac{-\sqrt{3}|x|}{2\sigma_x}}. \quad (5.5)$$

Более простой аппроксимацией является функция плотности вероятности Лапласа

$$p(x) = \frac{1}{\sqrt{2}\sigma_x} e^{\frac{-\sqrt{2}|x|}{\sigma_x}}. \quad (5.6)$$

На рис. 5.4 показана экспериментальная функция плотности вероятности совместно с функцией плотности вероятности Лапласа и

гамма-распределением. Все функции нормализованы таким образом, что среднее значение равно нулю, а дисперсия — единице. Хотя обе функции плотности вероятностей хорошо аппроксимируют экспериментальный результат, гамма-распределение, очевидно, обеспечивает лучшую аппроксимацию.

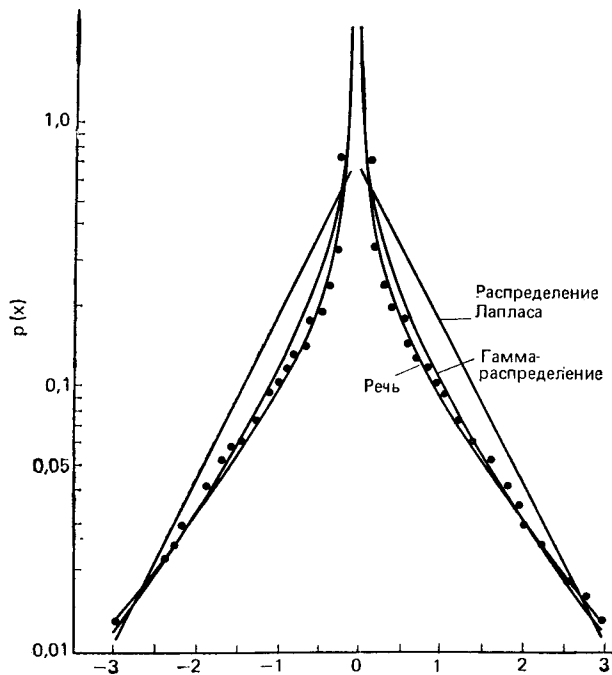


Рис. 5.4. Функции плотности вероятности речи, гамма-распределения и распределения Лапласа [3]

Автокорреляционная функция и спектральная плотность мощности речевого сигнала могут быть получены с использованием стандартных методов анализа временных рядов. Оценка автокорреляционной функции эргодического случайного процесса может быть получена путем усреднения за большой отрезок времени. Например, для получения усреднения за большой интервал времени достаточно немного изменить определение кратковременной автокорреляционной функции (4.30):

$$\hat{\phi}(m) = \frac{1}{L} \sum_{n=0}^{L-1-m} x(n)x(n+m), \quad 0 \leq |m| \leq L-1, \quad (5.7)$$

где L — большое целое число. Пример такой оценки показан на рис. 5.5 при частоте дискретизации 8 кГц [4]. Верхняя кривая вычислена по сигналу, пропущенному через фильтр нижних частот, нижняя — через полосовой фильтр. Заштрихованные области

вокруг каждой кривой показывают изменения в корреляции, возникающие для различных дикторов. Корреляция весьма велика между соседними отсчетами и быстро убывает при увеличении расстояния между ними. Видно, что речевой сигнал на выходе фильтра нижних частот более коррелирован, чем на выходе полосового фильтра.

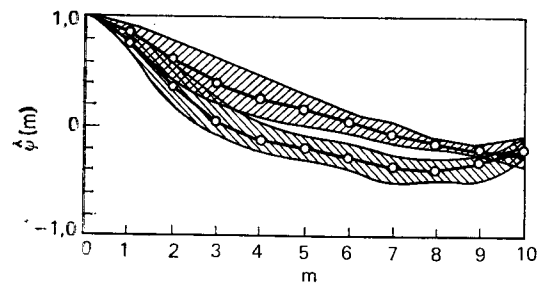


Рис. 5.5. Автокорреляционная функция речевых сигналов: верхняя кривая — для низкочастотной составляющей речи, нижняя — для высокочастотной [4]

Спектральную плотность мощности можно оценить различными путями. Для речевого сигнала один из наиболее ранних результатов был получен путем измерения сигнала на выходе гребенки полосовых фильтров [5]. На рис. 5.6 показан пример, в котором

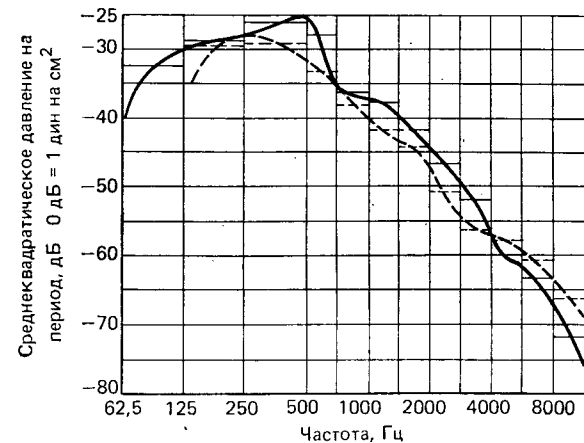


Рис. 5.6. Усредненная спектральная плотность мощности непрерывного речевого сигнала [5]: — шесть дикторов-мужчин; - - - пять дикторов-женщин

мощность усреднялась за минуту непрерывной речи. Этот рисунок показывает, что усредненная спектральная плотность мощности имеет максимум в диапазоне 250—500 Гц и затухает примерно на 8—10 дБ на октаву. Другой подход к оценке усредненной спектральной плотности состоит в оценивании $\hat{\phi}(m)$ соответственно (5.7) и последующем вычислении

$$\hat{\Phi}(e^{i\Omega T}) = \sum_{m=-M}^M w(m) \hat{\phi}(m) e^{-i\Omega m T} \quad (5.8)$$

для дискретной последовательности $\Omega_k = 2\pi k/T$ при $k=0, 1, \dots, \dots, N-1$, используя дискретное преобразование Фурье [6], где $w(m)$ окно (взвешивающая функция) для автокорреляционной функции. В качестве примера применения этого метода к оценке спектральной плотности речи на рис. 5.7 приведены соответствующие результаты при использовании окна Хемминга [7]. Еще один

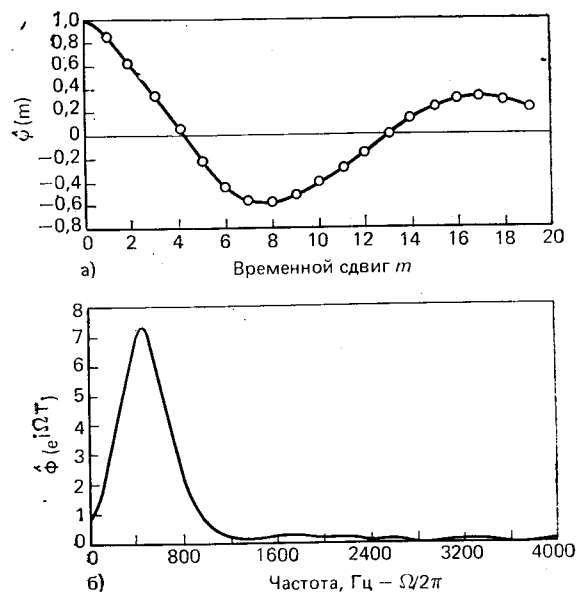


Рис. 5.7. Автокорреляционная функция (а) и спектральная плотность мощности (б) речевого сигнала [7]

подход состоит в вычислении передаточной функции цифрового фильтра, на входе которого действует белый шум, а сигнал на выходе имеет те же спектральные свойства, что и данный сигнал (см. гл. 8).

5.3. Квантование мгновенных значений

Как уже отмечалось, операции дискретизации и квантования удобно рассматривать отдельно, хотя часто разделить их затруднительно. Предположим, что речевой сигнал пропущен через фильтр нижних частот и в результате дискретизации получена последовательность непрерывных величин $\{x(n)\}$. В большинстве случаев в данной главе последовательность $\{x(n)\}$ рассматривается как случайный процесс в дискретном времени. Для того чтобы передать эту последовательность отсчетов по цифровому каналу связи, зарегистрировать ее в цифровом блоке памяти или использовать ее как входной сигнал некоторого алгоритма цифровой обработки,

каждый отсчет необходимо проквантовать до конечного множества значений, которые можно описать конечным множеством символов. Этот процесс квантования и кодирования изображен на рис. 5.8. Так же, как полезно разделять операции дискретизации и квантования, целесообразно разделить процесс представления последовательности $\{x(n)\}$ множеством символов на два этапа: квантование, результатом которого является последовательность величин $\{\hat{x}(n)\} = \{Q[x(n)]\}$, и кодирование, при котором каждой квантованной величине ставится в соответствие кодовое слово $c(n)$. Этот процесс изображен на рис. 5.8а. (Величина Δ на рисунке означает шаг квантования в квантователе.) Аналогично определим декодер как устройство, которое последовательности кодовых

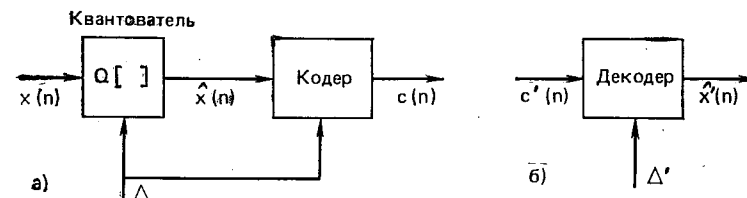


Рис. 5.8. Квантование и кодирование: а) кодер; б) декодер

слов $\{c'(n)\}$ ставит в соответствие последовательность квантованных отсчетов $\{\hat{x}'(n)\}$, как это показано на рис. 5.8б. Если последовательность кодовых слов $c'(n)$ точно совпадает с последовательностью кодовых слов $c(n)$, т. е. ошибки отсутствуют, то сигнал на выходе идеального декодера точно совпадает с последовательностью квантованных отсчетов входного сигнала, т. е. $\hat{x}'(n) = \hat{x}(n)$.

В большинстве случаев целесообразно для кодирования квантованных отсчетов использовать двоичную последовательность. С помощью B -разрядного двоичного кодового слова можно представить 2^B различных уровней квантования. Информационный объем цифрового представления, который надо знать при передаче или хранении сигнала, можно подсчитать:

$$I = B \cdot F_s = \text{скорость, бит/с}, \quad (5.9)$$

где F_s — частота дискретизации (т. е. отсч./с); B — число бит на отсчет сигнала. В общем случае желательно выбирать скорость передачи наиболее низкой, при которой еще сохраняется требуемое качество восприятия сигнала. Для данной полосы частот речевого сигнала минимальная частота дискретизации определяется теоремой о дискретизации. Таким образом, единственный путь уменьшения скорости передачи состоит в сокращении числа двоичных единиц на отсчет сигнала. С этой целью продолжим обсуждение различных способов квантования сигнала.

В общем случае целесообразно предполагать, что отсчеты сигнала будут попадать в конечный интервал значений, при котором

$$|x(n)| \leq X_{\max} \quad (5.10)$$

Для удобства следует предположить, что величина X_{max} бесконечно велика, что соответствует, например, функциям плотности вероятности гамма-распределения или Лапласа. Однако следует иметь в виду, что предположение о конечности диапазона значений в большей мере отвечает реальной ситуации. Даже если для описания сигнала используется функция плотности вероятности Лапласа, то легко показать (см. задачу 5.2), что только 0,35% отсчетов сигнала окажется вне диапазона

$$-4\sigma_x \leq x(n) \leq 4\sigma_x. \quad (5.11)$$

Таким образом, целесообразно считать, что полный размах сигнала пропорционален среднему квадратическому отклонению.

Диапазон изменения входного сигнала делится на интервалы, и операция квантования сводится к тому, что всем отсчетам входного сигнала, попавшим в некоторый интервал, присваивается одно

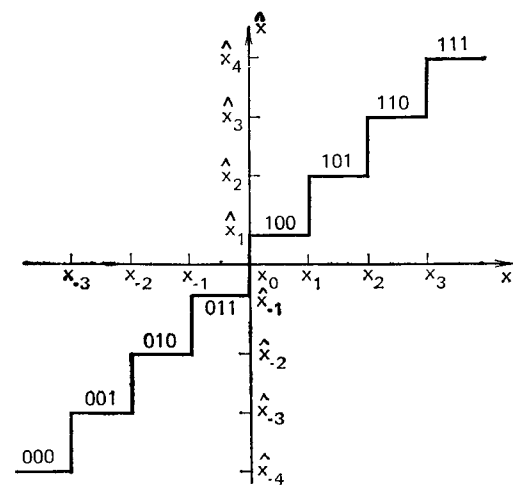


Рис. 5.9. Характеристика трехразрядного квантователя

и то же заданное значение. Этот процесс иллюстрирует рис. 5.9 для восьмиуровневого квантователя. Например, для всех значений входного сигнала $x(n)$, расположенных между x_1 и x_2 , значение сигнала на выходе будет $\hat{x}(n) = Q[x(n)] = \hat{x}_2$. Каждому уровню поставлено в соответствие трехразрядное слово, которым кодируется значение соответствующего уровня. Например, если отсчет попадает в интервал между x_1 и x_2 (рис. 5.9), то на выходе кодера появится слово 101. Конкретный способ кодирования уровней на рис. 5.9 произволен. В принципе можно использовать все восемь способов обозначения уровней, однако часто имеются причины выбирать вполне определенный способ кодирования.

5.3.1. Равномерное квантование

Интервалы и уровни квантования можно выбирать по-разному в зависимости от предлагаемого использования цифрового представления. Когда цифровое представление сигнала предназначено для обработки в некоторой системе, уровни и интервалы квантования выбирают обычно равномерно. Таким образом, для равномерного квантователя (см. рис. 5.9) получаем

$$x_i - x_{i-1} = \Delta \quad (5.12)$$

и

$$\hat{x}_i - \hat{x}_{i-1} = \Delta, \quad (5.13)$$

где Δ — шаг квантования. Для случая восьми уровней квантования на рис. 5.10 приведены характеристики двух обычно используемых квантователей. На рис. 5.10а изображен случай, когда на-

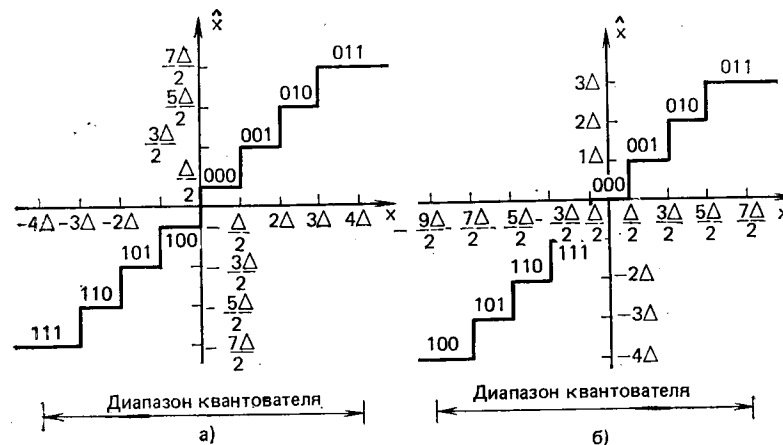


Рис. 5.10. Характеристики равномерных квантователей: а) с усечением; б) с округлением

чало отсчета приходится на середину вертикального участка ступенчатой функции. Этот класс квантователей называется квантователями с усечением. Аналогично на рис. 5.10б показан пример квантователей с округлением. Для случая, когда число уровней равно степени 2, как это обычно и бывает при использовании методов двоичного кодирования, квантователь с усечением имеет одинаковое количество положительных и отрицательных уровней, расположенных симметрично относительно начала координат. В отличие от этого квантователь с округлением имеет на один отрицательный уровень больше, но при этом обладает нулевым уровнем, который отсутствует у квантователя с усечением. Кодовые слова распределены по уровням на рис. 5.10 аналогично тому, как это сделано на рис. 5.9. Но в данном случае они обозначают непосредственно номер уровня в двоичной системе счисления. Например, если интерпретировать кодовые слова на рис. 5.10а как представление значения сигнала со знаком, полагая, что знаковый разряд — крайний слева, то уровни квантования связаны с кодовыми словами соотношением

$$\hat{x}(n) = (\Delta/2) \text{sign}(c(n)) + \Delta c(n), \quad (5.14)$$

где $\text{sign}(c(n)) = +1$, если первый разряд $c(n) = 0$, и $\text{sign}(c(n)) = -1$, если знаковый разряд $c(n) = 1$. Аналогично можно предста-

вить и кодовые слова на рис. 5.10б, но в этом случае соотношение между уровнями имеет вид

$$\hat{x}(n) = \Delta c(n). \quad (5.15)$$

Последний способ отображения уровней кодовыми словами используется обычно, когда последовательность отсчетов обрабатывается с представлением чисел в дополнительном коде (как в большинстве микропроцессоров), так как кодовые слова в этом случае являются просто численным значением отсчета сигнала.

Для описания равномерных квантователей (таких, как показанные на рис. 5.10) достаточно задать два параметра: число уровней и шаг квантования Δ . Число уровней выбирается обычно в виде 2^B с тем, чтобы использовать все B -разрядные кодовые слова. Параметры Δ и B выбираются таким образом, чтобы охватить весь диапазон сигнала. Если предположить, что $|x(n)| \leq X_{max}$, то (полагая симметричной функцию плотности вероятности $x(n)$) имеем

$$2X_{max} = \Delta \cdot 2^B. \quad (5.16)$$

При изучении эффектов квантования полезно представить квантованный сигнал в виде

$$\hat{x}(n) = x(n) + e(n), \quad (5.17)$$

где $x(n)$ — непрерывный отсчет, а $e(n)$ — ошибка, или шум квантования. Из рис. 5.10а и б легко установить, что если Δ и B выбрать в соответствии с (5.16), то

$$-\Delta/2 \leq e(n) \leq \Delta/2. \quad (5.18)$$

Если выбирать в качестве примера размах сигнала в $8\sigma_x$ и предположить, что сигнал имеет распределение Лапласа, то только 0,35% отсчетов окажутся вне диапазона квантователя. Квантование этих отсчетов будет сопровождаться ошибкой, большей $\pm \Delta/2$, то их число крайне мало, поэтому целесообразно выбирать диапазон квантования около $8\sigma_x$ и пренебречь при теоретическом анализе редкими большими ошибками [8].

Очевидно, что нам известен только $\hat{x}(n)$, а $x(n)$ и $e(n)$ неизвестны. Для изучения эффектов квантования удобно и полезно предположить простую статистическую модель шума квантования. Эта модель основана на следующих предположениях.

1. Шум квантования является стационарным белым шумом, т. е.

$$E[x(n)e(n+m)] = \begin{cases} \sigma_e^2, & m=0, \\ 0, & \text{в противном случае} \end{cases} \quad (5.19)$$

2. Шум квантования не коррелирован с входным сигналом, т. е.

$$E[x(n)e(n+m)] = 0, \text{ для всех } m \quad (5.20)$$

3. Распределение шума равномерно в любом интервале квантования и поскольку все интервалы равны между собой, то

$$p_e(e) = \begin{cases} \frac{1}{\Delta}, & -\frac{\Delta}{2} \leq e \leq \frac{\Delta}{2}, \\ 0, & \text{в противном случае} \end{cases} \quad (5.21)$$

Очевидно, что эти предположения не выполняются для некоторых сигналов. Они нарушаются, например, если входной сигнал постоянен для всех n . Речевой сигнал, однако, является достаточно сложным и быстроизменяющимся в пределах любого уровня квантования, и если шаг Δ достаточно мал, то вероятность попадания двух последовательных отсчетов в различные далеко отстоящие интервалы достаточно велика. Справедливость сделанных выше предположений подтверждается экспериментом [9].

Пример, иллюстрирующий правомерность предположений, приведен на рис. 5.11 [6]. На рис. 5.11а изображено 400 последовательных отсчетов сигнала, которые были подвергнуты операции

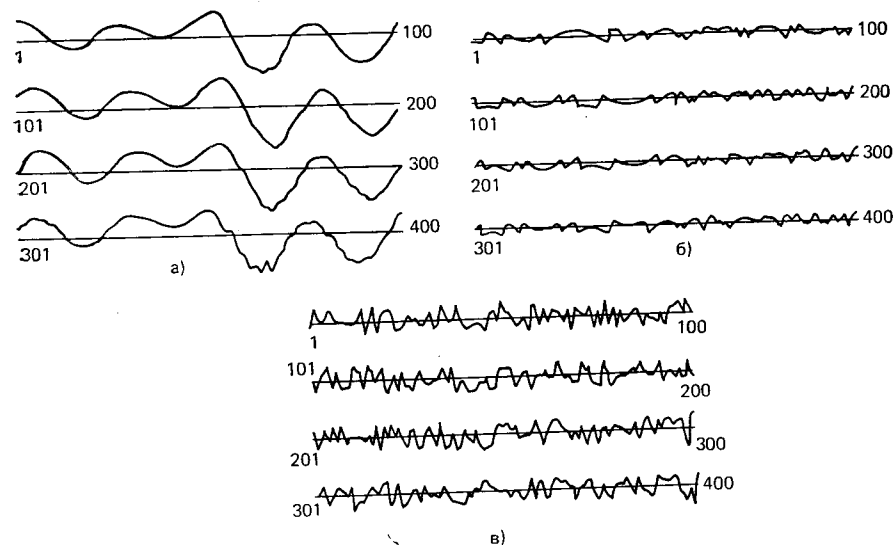


Рис. 5.11. Временная диаграмма речи (а); шум квантования для трехразрядного квантователя (б); шум квантования для восьмиразрядного квантователя (в) (увеличено в 66 раз по сравнению с (а) и (б))

квантования в трех- и восьмиразрядном квантователях¹ (рис. 5.10б). Ошибки квантования изображены на рис. 5.11б и в соответственно. В случае трехразрядного квантования видна корреляция сигнала и ошибки, в то время как во втором случае такая корреляция не наблюдается. Для подтверждения этого факта на

¹ Термин восьмиразрядный (8-bit) квантователь означает 2⁸-уровневый квантователь. (Прим. ред.)

рис. 5.12а и в представлены корреляционные функции для первого и второго случаев соответственно. Очевидно, что корреляционная функция рис. 5.12в более согласуется с предложением, что $\varphi(m) = \sigma_e^2 \delta(m)$, поскольку на рис. 5.12а видна значительная корреляция при $m > 0$. Этот же эффект наблюдается и на спектральной плотности мощности (рис. 5.12б и г). Спектр погрешности

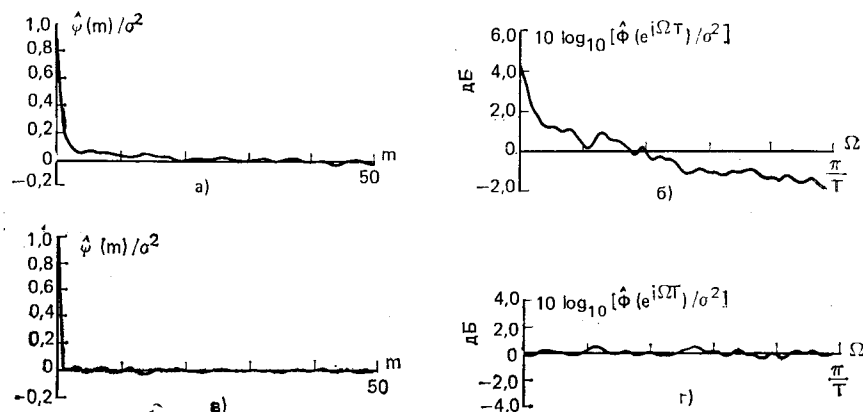


Рис. 5.12. Нормализованная оценка автокорреляции для трехразрядного квантователя (а); спектр мощности для трехразрядного квантователя (б), нормализованная оценка автокорреляции для восьмиразрядного квантователя (в); спектр мощности для восьмиразрядного квантователя (г)

при трехразрядном квантовании уменьшается в области высших частот (как и спектр речи), а для восьмиразрядного квантователя спектр равномерен. Заметим однако, что даже для трехразрядного квантователя спектр затухает со скоростью 6 дБ на октаву.

В рамках введенной статистической модели можно связать мощность шума и сигнала с параметрами квантователя. Для этой цели удобно вычислить отношение сигнал/шум квантования, определяемое выражением¹

$$SNR = \frac{\sigma_x^2}{\sigma_e^2} = \frac{E[x^2(n)]}{E[e^2(n)]} = \frac{\sum_n x^2(n)}{\sum_n e^2(n)}. \quad (5.22)$$

Если предположить, что диапазон непрерывных значений равен $2X_{max}$, то для B -разрядного квантователя получаем

$$\Delta = 2X_{max}/2^B. \quad (5.23)$$

Если предположить равномерное распределение шума, получим (см. задачу 5.1)

$$\sigma_e^2 = \Delta^2/12 = X_{max}^2/(3)2^{2B}. \quad (5.24)$$

¹ Предполагается, что сигнал имеет нулевое среднее значение. В противном случае оно вычитается из сигнала перед вычислением отношения сигнал/шум.

Подставляя (5.24) в (5.22), имеем

$$SNR = (3)2^{2B}/[X_{max}/\sigma_x]^2 \quad (5.25)$$

или, выражая отношение сигнал/шум в децибелах,

$$SNR = 10 \log_{10} [\sigma_x^2/\sigma_e^2] = 6B + 4,77 - 20 \log_{10} [X_{max}/\sigma_x]. \quad (5.26)$$

Предполагая диапазон квантования $X_{max} = 4\sigma_x$, из (5.26) получим

$$SNR = 6B - 7,2. \quad (5.27)$$

Соотношение (5.27), из которого следует, что каждое добавление одного разряда в кодовом слове улучшает отношение сигнал/шум на 6 дБ, справедливо при следующих предположениях: 1) входной сигнал изменяется таким образом, что справедлива описанная ранее статистическая модель шума квантования; 2) шаг квантования мал настолько, что шум белый и не коррелирован с сигналом; 3) диапазон квантования установлен таким образом, что он превышает размах сигнала. Следовательно, диапазон квантования используется полностью, и в тоже время количество отсчетов, не попадающих в него, достаточно мало.

Для речевых сигналов первые два предположения выполняются, если количество уровней квантования больше, чем 2^6 . Однако третье предположение менее справедливо, поскольку энергия сигнала может изменяться более чем на 40 дБ в зависимости от диктора и условий передачи. Даже для данного диктора амплитуда речевого сигнала существенно меняется при переходе от вокализованной речи к невокализованной и на протяжении вокализованных сегментов. Соотношение (5.27) предполагает полное использование диапазона квантования, и если размах сигнала очень мал, то это эквивалентно использованию лишь нескольких уровней квантования квантователя. Например, из (5.26) следует, что в случае если дисперсия входного сигнала составляет лишь половину той, на которую рассчитан квантователь, то это приведет к ухудшению отношения сигнал/шум на 6 дБ. В то же время из кратковременного анализа сигнала известно, что дисперсия на невокализованных сегментах может быть на 20—30 дБ меньше, чем дисперсия на вокализованных, т. е. кратковременное отношение сигнал/шум на невокализованных сегментах может быть значительно меньше, чем на вокализованных.

Для поддержания ошибки квантования на приемлемом уровне необходимо выбирать значительно больше уровней квантования, чем это следует из предварительного анализа в рамках предположения о стационарности сигнала. Например, использование соотношения (5.27) позволяет сделать вывод, что значение $B=7$ обеспечивает отношение сигнал/шум, равное 36 дБ, т. е. хорошее качество связи. Однако известно, что необходимо около 11 разрядов квантователя для получения высококачественного речевого сигнала при равномерном квантовании.

Таким образом, желательно иметь устройство квантования, при котором отношение сигнал/шум не зависит от уровня сигнала,

т. е. вместо постоянной не зависящей от уровня сигнала ошибки (как это имеет место при равномерном квантовании) хотелось бы получить постоянную относительную ошибку. Это достигается путем использования неравномерного распределения уровней квантования.

5.3.2. Мгновенное компандирование

Для того чтобы относительная ошибка была постоянной, уровни квантования должны быть распределены логарифмически. С другой стороны, вместо квантования исходного сигнала для достижения постоянной ошибки можно квантовать его логарифм. Этот процесс изображен на рис. 5.13, где входной сигнал компрес-

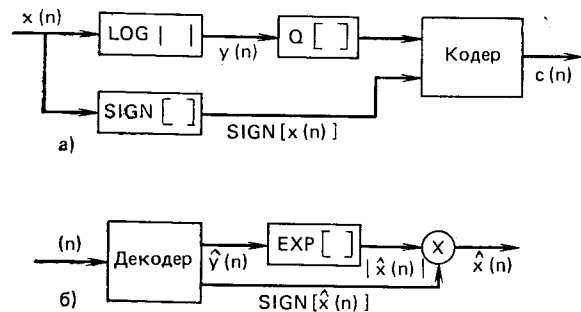


Рис. 5.13. Структурная схема системы логарифмического кодирования

сируется перед квантованием с помощью логарифмического преобразования, а выходной сигнал после декодирования экспандируется с помощью экспоненциального преобразования. Убедимся, что это приводит к требуемой нечувствительности ошибки квантования от значения сигнала. Предположим, что

$$y(n) = \ln |x(n)|. \quad (5.28)$$

Обратное преобразование равно

$$x(n) = \exp [y(n)] \operatorname{sign} [x(n)], \quad (5.29)$$

где $\operatorname{sign} [x(n)] = +1$, если $x(n)$ положительно, и $\operatorname{sign} [x(n)] = -1$, если $x(n)$ отрицательно. Теперь квантованный логарифм имеет вид

$$\hat{y}(n) = Q [\log |x(n)|] = \log |x(n)| + \varepsilon(n), \quad (5.30)$$

где, как и ранее, предполагается, что $\varepsilon(n)$ не зависит от $\log |x(n)|$. Применяя к квантованной величине обратное преобразование, получаем

$$\begin{aligned} \hat{x}(n) &= \exp [\hat{y}(n)] \operatorname{sign} [x(n)] = |x(n)| \operatorname{sign} [x(n)] \exp [\varepsilon(n)] = \\ &= x(n) \exp [\varepsilon(n)]. \end{aligned} \quad (5.31)$$

Если $\varepsilon(n)$ мало, можно аппроксимировать это уравнение в виде

$$\hat{x}(n) \approx x(n) [1 + \varepsilon(n)] = x(n) + \varepsilon(n) x(n) = x(n) + f(n), \quad (5.32)$$

где $f(n) = x(n)\varepsilon(n)$. Таким образом, поскольку $x(n)$ и $\varepsilon(n)$ предполагаются некоррелированными, получаем

$$\sigma_f^2 = \sigma_x^2 \sigma_\varepsilon^2; \quad SNR = \sigma_x^2 / \sigma_f^2 = 1 / \sigma_\varepsilon^2. \quad (5.33); (5.34)$$

Следовательно, отношение сигнал/шум не зависит от мощности сигнала, а зависит только от шага квантования. Квантователь такого типа не имеет практического значения, поскольку динамический диапазон (отношение максимального значения к минимальному) бесконечен и, таким образом, требуется бесконечное число уровней квантования. Выполненный анализ, возможно и лишенный практического смысла, позволяет, однако, сделать вывод о том, что характеристика компрессора может быть близкой к логарифмической. Использование системы компрессор—экспандер для

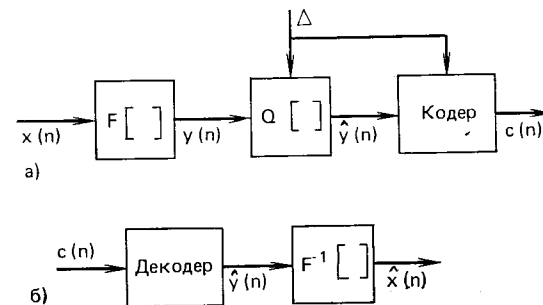


Рис. 5.14. Структурная схема системы компрессор-экспандер для квантования

квантования показано на рис. 5.14. Смит [10] исследовал характеристики компрессора, названные μ -законом компандирования. В этом случае

$$y(n) = F [x(n)] = X_{max} \frac{\log \left[1 + \mu \frac{|x(n)|}{X_{max}} \right]}{\log [1 + \mu]} \operatorname{sign} [x(n)]. \quad (5.35)$$

На рис. 5.15 представлено семейство зависимостей $y(n)$ от $x(n)$ для различных значений μ . Очевидно, что использование функции (5.35) решает проблему малых амплитуд, поскольку $y(n) = 0$, если $|x(n)| = 0$. При $\mu = 0$ уравнение (5.35) превращается в равенство

$$y(n) = x(n), \quad (5.36)$$

т. е. уровни квантования располагаются равномерно. Однако для больших μ и больших $|x(n)|$ получаем

$$|y(n)| \approx X_{max} \log |x(n)/X_{max}|. \quad (5.37)$$

Таким образом, за исключением очень малых амплитуд, кривые, соответствующие μ -закону, позволяют получить постоянный процент дисперсии шума от дисперсии сигнала. На рис. 5.16 показано

распределение уровней квантования при $\mu=40$ и восьми уровнях (характеристика квантователя антисимметрична относительно начала координат).

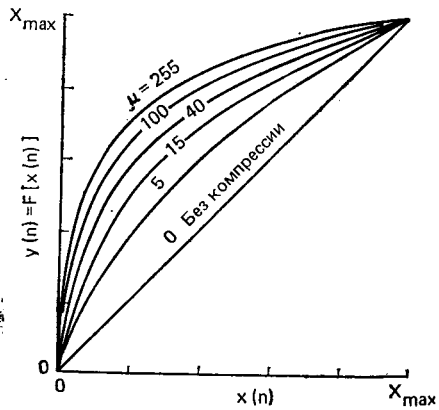


Рис. 5.15. Характеристика компрессии по μ -закону [10]

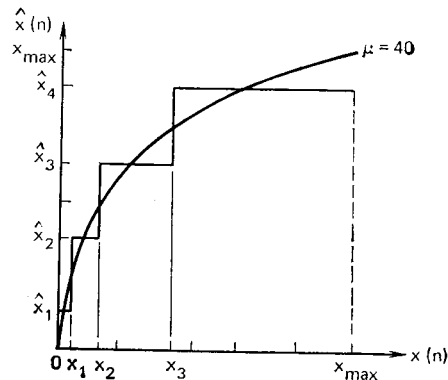


Рис. 5.16. Распределение уровней квантования для μ -закона в трехразрядном квантователе при $\mu=40$

Используя предположения такого же типа, как и в случае анализа равномерного квантователя, Смит [10] получил формулу для отношения сигнал/шум при использовании μ -закона в виде

$$SNR = 6B + 4,77 - 20 \log_{10} [\ln(1 + \mu)] - 10 \log_{10} \left[1 + \left(\frac{X_{max}}{\mu\sigma_x} \right)^2 + \sqrt{2} \left(\frac{X_{max}}{\mu\sigma_x} \right) \right], \quad (5.38)$$

Сравнивая это уравнение с (5.26), можно отметить, что в данном случае отношение сигнал/шум значительно меньше зависит от величины (X_{max}/σ_x) . Эта зависимость уменьшается при возрастании μ , т. е. хотя член $20 \log_{10} [\ln(1 + \mu)]$ уменьшает отношение сигнал/шум, второе слагаемое с ростом μ возрастает. На рис. 5.17 и 5.18 графически представлены соотношения (5.26) и (5.38) как функции величины X_{max}/σ_x при $\mu=100$ и 500 соответственно. Величина X_{max} является параметром устройства квантования. Она определяет порог «переполнения», т. е. значение, выше которого все отсчеты ограничиваются. Величина σ_x является параметром сигнала, определяющим «среднюю» амплитуду сигнала. Величина X_{max}/σ_x показывает, насколько диапазон сигнала согласован с диапазоном квантователя. Пунктирные кривые на рис. 5.17 иллюстрируют зависимость отношения сигнал/шум в децибелах от X_{max}/σ_x . При заданном значении X_{max} уменьшение вдвое величины σ_x приводит к потере в отношении сигнал/шум на 6 дБ. Для заданного значения X_{max}/σ_x отношение сигнал/шум возрастает на 6 дБ при добавлении одного разряда квантователя. Это справед-

ливо как для равномерного квантователя, так и при использовании μ -закона.

Поясним правомерность соотношений (5.26) и (5.38) и кривых рис. 5.17 и 5.18. Одно из предположений, сделанных при выводе этих уравнений, состоит в том, что количество переполнений в квантователе пренебрежимо мало, т. е. вероятность того, что некоторый отсчет превысит X_{max} , пренебрежимо мала. Это предположение, очевидно, нарушается, если дисперсия σ_x приблизительно равна X_{max} , т. е. $X_{max}/\sigma_x \approx 1$. Экспериментальные кривые отношения сигнал/шум подтверждают сделанное утверждение. Соотношения (5.26) и (5.38) хорошо описывают отношение сигнал/шум при $(X_{max}/\sigma_x) > 8$ [10].

Важное свойство μ -закона, иллюстрируемое этими кривыми, состоит в том, что отношение сигнал/шум более или менее постоянно в широком диапазоне. Например, из рис. 5.17 видно, что при $\mu=100$ отношение сигнал/шум уменьшается всего на 2 дБ при

$$8 < X_{max}/\sigma_x < 30, \quad (5.39)$$

а из рис. 5.18 видно, что при $\mu=500$ отношение сигнал/шум — менее чем на 2 дБ в диапазоне

$$8 < X_{max}/\sigma_x < 150. \quad (5.40)$$

Однако сравнение рис. 5.17 и 5.18 показывает, что максимальное отношение сигнал/шум во втором случае уменьшается на 2,6 дБ. Таким образом, используя большие значения коэффициента компрессии, мы получаем выигрыш в динамическом диапазоне ценой проигрыша в отношении сигнал/шум.

Как следует из рис. 5.17 и 5.18, при $B=7$ отношение сигнал/шум, равное 34 дБ, достигается в широком диапазоне уровней

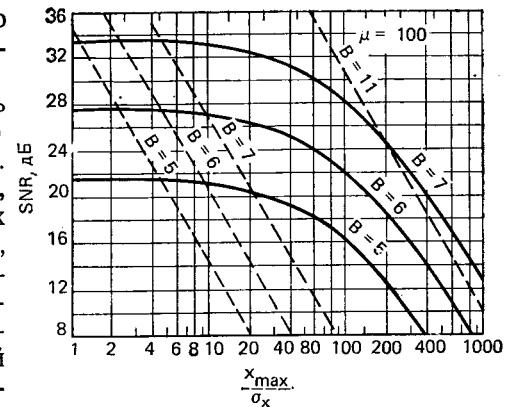


Рис. 5.17. SNR для μ -закона и равномерного квантователя как функция X_{max}/σ_x при $\mu=100$ и различном числе разрядов B [10]

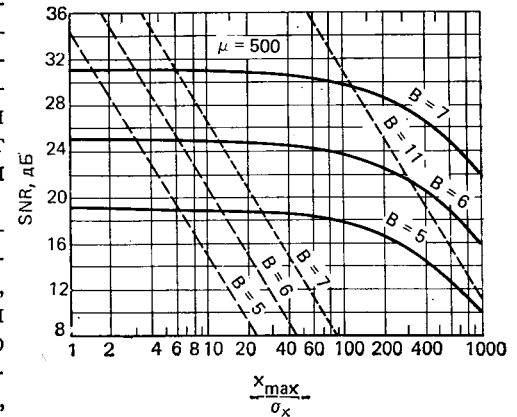


Рис. 5.18. SNR для μ -закона и равномерного квантователя при $\mu=500$, $B=5$; 6; 7; 11 [10]

входного сигнала. Поэтому семirazрядная ИКМ с компрессией используется как стандарт для получения речевого сигнала с хорошим качеством. При равномерном квантовании для получения такого же динамического диапазона требуется 11 разрядов. Как следует из рис. 5.18, 11-разрядное равномерное квантование оказывается лучше семirazрядного квантования для $\mu=500$ ($\sigma_v > > 0,01X_{max}$). Таким образом, можно сказать, что 11-разрядное равномерное квантование будет таким же или лучшим, чем семirazрядное квантование при $\mu=500$ для уровней входного сигнала, составляющих, по крайней мере, 1% максимального уровня квантования.

5.3.3. Оптимальное квантование

Квантование по μ -закону позволяет получить постоянное отношение сигнал/шум в широком диапазоне дисперсий входного сигнала. Как отмечалось выше, это достигается ценой ухудшения отношения сигнал/шум по сравнению с тем, которое можно получить, если диапазон квантования согласован с дисперсией сигнала. В тех случаях, когда дисперсия сигнала известна, можно так выбирать уровни квантования, чтобы минимизировать мощность шума и, таким образом, максимизировать отношение сигнал/шум квантования. Эта задача изучалась Максом [11] и позже Паезом и Глиссоном [3]. Дисперсия шума квантования имеет вид

$$\sigma_e^2 = E[e^2(n)] = E[(\hat{x}(n) - x(n))^2], \quad (5.41)$$

где $\hat{x}(n) = Q[x(n)]$. На основе рис. 5.9 предположим, что имеется M уровней квантования, которые можно обозначить через $\{\hat{x}_{-M/2}, \hat{x}_{-M/2+1}, \dots, \hat{x}_{-1}, \hat{x}_1, \dots, \hat{x}_{M/2}\}$, полагая, что M четное. Уровень квантования, соответствующий интервалу от x_{j-1} до x_j , обозначен как \hat{x}_j . Для симметричной функции плотности вероятности с нулевым средним удобно обозначить центральную граничную точку $x_0 = 0$, и если функция плотности вероятности не обращается в нуль при больших значениях, как, например, функция Лапласа, то максимальное значение границ внешнего интервала квантования будет $\pm \infty$, т. е. $x_{\pm M/2} = \pm \infty$. В этих предположениях можно записать

$$\sigma_e^2 = \int e^2 p_e(e) de. \quad (5.42)$$

На рис. 5.19 показана зависимость e от x . Как видно из рисунка, вклад каждого интервала квантования в функцию плотности вероятности погрешности оценить сложно. Поскольку

$$e = \hat{x} - x, \quad (5.43)$$

Можно выполнить замену переменной в (5.42)

$$p_e(e) = p_e(\hat{x} - x) = p_{x/\hat{x}}(\hat{x}/x) \Delta_x p_x(x), \quad (5.44)$$

что приводит к соотношению

$$\sigma_e^2 = \sum_{i=-M/2+1}^{M/2} \int_{x_{i-1}}^{x_i} (\hat{x}_i - x)^2 p(x) dx \quad (5.45)$$

(отметим, что в данную формулу для дисперсии шума входят и ошибки за счет ограничения или переполнения). Если $p(x) =$

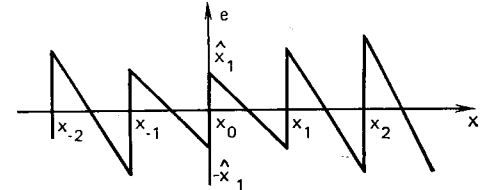


Рис. 5.19. Ошибка квантования как функция уровня x неравномерно квантователя

$= p(-x)$, то характеристика квантования будет антисимметричной, т. е. $\hat{x}_i = -\hat{x}_{-i}$ и $x_i = -x_{-i}$. Таким образом,

$$\sigma_e^2 = 2 \sum_{i=1}^{M/2} \int_{x_{i-1}}^{x_i} (\hat{x}_i - x)^2 p(x) dx. \quad (5.46)$$

Теперь требуется выбрать множество таких параметров $\{x_i\}$ и $\{\hat{x}_i\}$, чтобы минимизировать σ_e^2 . Для решения этой задачи продифференцируем σ_e^2 по этим параметрам и приравняем производные к нулю. Это приводит к системе уравнений

$$\int_{x_{i-1}}^{x_i} (\hat{x}_i - x) p(x) dx = 0, \quad i = 1, 2, \dots, M/2; \quad (5.47a)$$

$$x_i = (\hat{x}_i + \hat{x}_{i+1})/2, \quad i = 1, 2, \dots, (M/2) - 1 \quad (5.47b)$$

в предположении

$$x_0 = 0; \quad (5.48a)$$

$$x_{\pm M/2} = \pm \infty. \quad (5.48b)$$

Уравнения (5.47) показывают, что оптимальные пороги равны полусумме соответствующих уровней квантования. Уравнение (5.47a) означает, что соответствующие уровни есть средние значения функции плотности вероятности на интервале от x_{i-1} до x_i . Два этих множества уравнений следует решать совместно относительно $M-1$ неизвестных параметров квантователя. Поскольку данные уравнения нелинейны, замкнутое решение можно получить лишь в специальных случаях. Для других случаев следует использовать итеративные методы решения. Такой метод последовательных приближений дан Максом [1]. Паез и Глиссон использовали эту процедуру для получения оптимальных порогов в случае распределения Лапласа и гамма-распределения [3].

В общем случае решение уравнений (5.47) приводит к неравномерному распределению уровней квантования. Только при равномерном распределении сигнала оптимальное решение будет равномерным, т. е.

$$\hat{x}_i - \hat{x}_{i-1} = x_i - x_{i-1} = \Delta. \quad (5.49)$$

Полагая квантователь равномерным, можно определить значение оптимального шага квантования Δ , при котором минимальна мощность шума квантования и максимально отношение сигнал/(шум квантования). В этом случае

$$x_i = \Delta \cdot i; \quad (5.50)$$

$$\hat{x}_i = (2i-1) \Delta/2 \quad (5.51)$$

и Δ удовлетворяет уравнению

$$\sum_{i=1}^{M/2-1} (2i-1) \int_{(i-1)\Delta}^{i\Delta} \left[\left(\frac{2i-1}{2} \right) \Delta - x \right] p(x) dx + (M-1) \int_{(M/2-1)\Delta}^{\infty} \left[\left(\frac{M-1}{2} \right) \Delta - x \right] p(x) dx = 0. \quad (5.52)$$

Таблица 5.1

Оптимальный квантователь для сигналов с распределением Лапласа ($m_x=0, \sigma_x=1$) [3]

N	2		4		8		16		32	
	x_j	\hat{x}_i	x_i	\hat{x}_i	x_i	\hat{x}_i	x_i	\hat{x}_i	x_i	\hat{x}_i
1	∞	0,707	1,102	0,395	0,504	0,222	0,266	0,126	0,147	0,072
2			∞	1,810	1,181	0,785	0,566	0,407	0,302	0,222
3					2,285	0,910	0,726	0,467	0,382	
4					∞	2,994	1,317	1,095	0,642	0,551
5							1,821	1,540	0,829	0,732
6							2,499	2,103	1,031	0,926
7							3,605	2,895	1,250	1,136
8							∞	4,316	1,490	1,365
9									1,756	1,616
10									2,055	1,896
11									2,398	2,214
12									2,804	2,583
13									3,305	3,025
14									3,978	3,586
15									5,069	4,371
16									∞	5,768
MSE	0,5		0,1765		0,0548		0,0154		0,00414	
SNR, дБ	3,01		7,53		12,61		18,12		23,83	

Если $p(x)$ известна или задана (например, функция плотности вероятности Лапласа), то интегралы могут быть вычислены. Тогда получается простое уравнение, которое можно решить на ЭВМ, пользуясь итеративными методами и изменяя Δ до получения оптимального значения.

В табл. 5.1 и 5.2 содержатся оптимальные уровни квантования для гамма-распределения и распределения Лапласа [3]. (Результаты получены в предположении единичной дисперсии. Если в действительности дисперсия σ_x^2 , то каждое число в таблице следует

Таблица 5.2

Оптимальный квантователь для сигналов с гамма-распределением ($m_x=0, \sigma_x^2=1$) [3]

N	2		4		8		16		32	
	x_j	\hat{x}_i	x_i	\hat{x}_i	x_i	\hat{x}_i	x_i	\hat{x}_i	x_i	\hat{x}_i
1	∞	0,577	1,205	0,302	0,504	0,149	0,229	0,072	0,101	0,033
2			∞	2,108	1,401	0,859	0,588	0,386	2,252	0,169
3					2,872	1,944	1,045	0,791	0,429	0,334
4					∞	3,799	1,623	1,300	0,630	0,523
5							2,372	1,945	0,857	0,737
6							3,407	3,798	1,111	0,976
7							5,050	4,015	1,397	1,245
8							∞	6,085	1,720	1,548
9									2,089	1,892
10									2,517	2,287
11									3,022	2,747
12									3,633	3,296
13									4,404	3,970
14									5,444	4,838
15									7,046	6,050
16									∞	8,043
MSE	0,6680		0,2326		0,0712		0,0196		0,0052	
SNR, дБ	1,77		6,33		11,47		17,07		22,83	

умножить на σ_x .) На рис. 5.20 изображена схема трехразрядного (восьмиуровневого) квантователя для функции плотности вероятности Лапласа. Легко видеть, что шаг квантования возрастает при убывании значения функции плотности вероятности. Это согласуется с интуитивным представлением о том, что большие значения ошибки квантования должны соответствовать наиболее редко возникающим отсчетам. Сравнение рис. 5.16 и 5.20 позволяет отметить сходство оптимального квантователя с квантователем на основе μ -закона. Таким образом, можно ожидать, что неравномерный квантователь обладает улучшенным динамическим диапазоном. Это подтверждается результатами работы [3].

На рис. 5.21 показана зависимость шага квантования оптимального равномерного квантователя для гамма-распределения,

распределений Лапласа [3] и Гаусса [11] от числа уровней. Как это и можно было ожидать, шаг квантования убывает приблизительно экспоненциально с увеличением разрядности кодовых слов. Некоторые различия в форме кривых объясняются различиями в

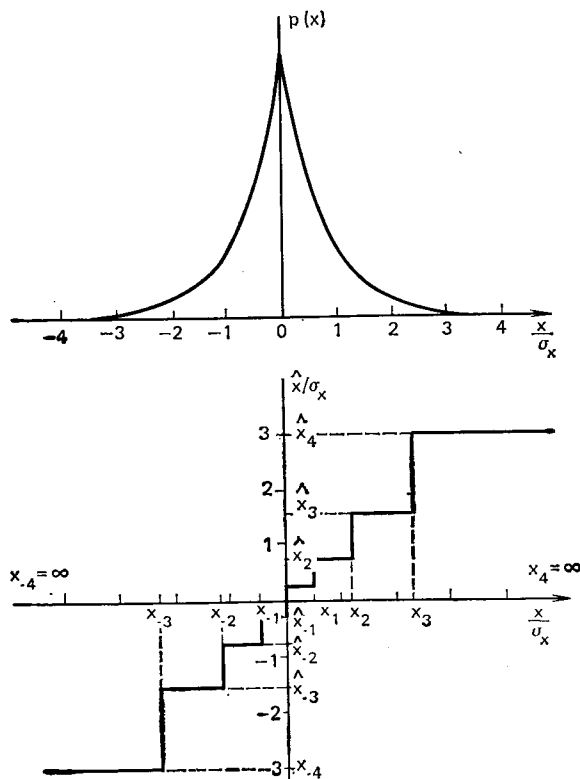


Рис. 5.20. Функция плотности вероятности и характеристика квантователя для распределения Лапласа и неравномерного квантователя с тремя разрядами

виде функции плотности вероятности. Хотя оптимальное квантование дает минимальные погрешности при точном соответствии дисперсии и распределении сигнала, нестационарный характер речевого сигнала в системах связи приводит к неудовлетворительным результатам и в этом случае. Простым подтверждением этого обстоятельства является ситуация, возникающая в системах передачи в период молчания, т. е. при занятом канале передачи. При этом сигнал на входе квантователя весьма мал (предполагается малый шум), что приводит к колебаниям сигнала на выходе квантователя между наименьшими уровнями квантования.

Для симметричного квантователя (см. рис. 5.10а), если внутренние уровни квантования больше, чем мгновенные значения шума, мощность шума на выходе будет больше мощности шума на

входе. Поэтому применение оптимального квантователя с малым числом уровней квантования не имеет практического смысла. В табл. 5.3 для сравнения представлены значения наименьшего уровня квантования [3] для ряда оптимальных равномерных и неравномерных квантователей и квантователя, построенного по μ -закону при $\mu=100$. Легко видеть, что квантователь, построенный по μ -закону, приводит к меньшему шуму незанятого канала по сравнению с любым оптимальным квантователем. А для больших значений μ наименьший уровень квантования лежит еще ниже (так при $\mu=255$ минимальный уровень составляет 0,031). Поэтому μ -квантователь чаще используется на практике, несмотря на несколько меньшее отношение сигнал/шум по сравнению с оптимальным.

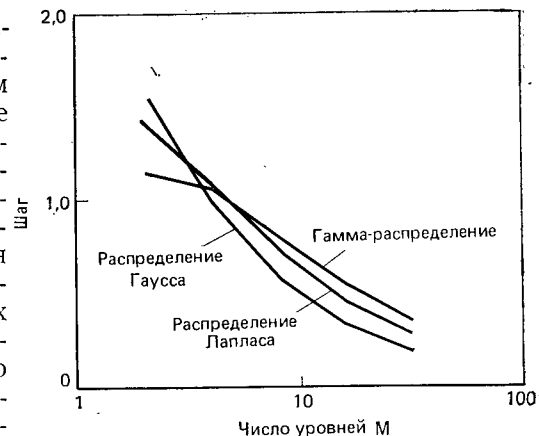


Рис. 5.21. Оптимальные размеры шага равномерного квантователя при лапласовском, гамма- и гауссовском распределениях [11]

Таблица 5.3
Отношение сигнал/шум для восьмиуровневого квантователя [12]

Закон	SNR, дБ	Наименьший уровень ($\delta_x=1$)
μ -закон ($x_{max}=8\sigma_x=\mu=100$)	9,5	0,062
Гаусса	14,6/14,3*	0,245/0,293
Лапласа	12,6/11,4	0,222/0,366
Гамма	11,5/11,5	0,149/0,398
Речь	12,1/8,4	0,124/0,398

* В числителе — для неравномерного, в знаменателе — для равномерного квантователя.

5.4. Адаптивное квантование

Как следует из результатов предыдущего параграфа, при квантовании речевого сигнала возникают серьезные трудности. С одной стороны, шаг квантования следует выбирать достаточно большим для согласования диапазона квантования с размахом сигнала. С другой стороны, шаг квантования следует сделать малым для уменьшения шума квантования. Это еще более усложняется нестационарным характером речевого сигнала и его последующей передачей по каналу связи. Амплитуда речевого сигнала может

изменяться в широких пределах в зависимости от диктора, условий передачи, а также внутри фразы при переходе от вокализованного к невокализованному сегменту. Как уже отмечалось, один из методов учета этих флуктуаций состоит в применении неравномерного квантования. Другой метод состоит в адаптации свойств квантователя к уровню входного сигнала. В этом параграфе обсуждаются основные принципы адаптивного квантования, а в последующих приводятся примеры применения методов адаптивного квантования совместно с линейным предсказанием речи. Если адаптивное квантование применяется непосредственно к отсчетам входного сигнала, то такой метод обработки называется адаптивной ИКМ или сокращенно АИКМ.

Основная идея адаптивного квантования состоит в том, что шаг квантования (или, в общем случае, интервалы и уровни квантования) изменяется таким образом, чтобы соответствовать изменяющейся дисперсии входного сигнала (рис. 5.22а). Другой метод адаптации (рис. 5.22б) соответствует случаю, когда характе-

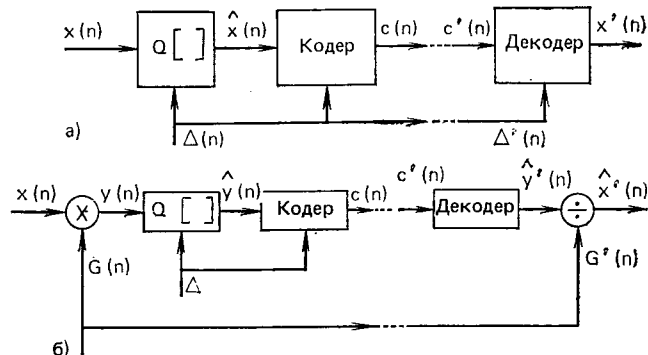


Рис. 5.22. Структурная схема адаптивного квантователя; а) с переменным шагом; б) с переменным усилением

ристики квантователя не изменяются, а постоянный уровень дисперсии поддерживается за счет переменного коэффициента усиления. В первом случае шаг квантования должен увеличиваться или уменьшаться при увеличении или уменьшении дисперсии входного сигнала соответственно. В случае неравномерного квантования это приводит к соответствующему масштабированию интервалов и уровней квантования. Во втором подходе, применимом в равной мере как к равномерному, так и к неравномерному квантователям, коэффициент усиления изменяется обратно пропорционально дисперсии входного сигнала так, чтобы поддерживать ее постоянной. В обоих случаях необходимо оценивать изменяющиеся во времени характеристики сигнала.

При изучении текущих свойств сигнала полезно выяснить, как быстро происходят их изменения. Значения сигнала изменяются от отсчета к отсчету или на малом количестве отсчетов и могут

рассматриваться как *мгновенные*. Максимальное значение амплитуды сигнала на невокализованном или вокализованном сегменте речи остаются относительно постоянными в течение длительного интервала времени. Такие изменения называются *слоговыми*, и это означает, что они проявляются на интервалах времени, сравнимых по протяженности с длительностью одного слога. При рассмотрении методов квантования целесообразно их классифицировать в соответствии с тем, медленно или быстро происходит адаптация, т. е. является она слоговой или мгновенной.

Имеется два класса схем адаптивного квантования. В первом: амплитуда или дисперсия входного сигнала оценивается непосредственно по этому сигналу. Такие схемы называются квантователями с адаптацией по входу. В схемах второго класса шаг квантования подстраивается по выходному сигналу $\hat{x}(n)$ или, что то же самое, по выходной последовательности кодовых слов $c(n)$. Это квантователи с адаптацией по *выходу*. В обоих случаях адаптация может быть как слоговой, так и мгновенной.

5.4.1. Адаптация по входному сигналу

На рис. 5.23 показана в общем виде схема квантователя с адаптацией по входному сигналу. Для простоты будем полагать, что квантователь равномерный и, таким образом, достаточно изменить только шаг квантования. Полученные результаты затем:

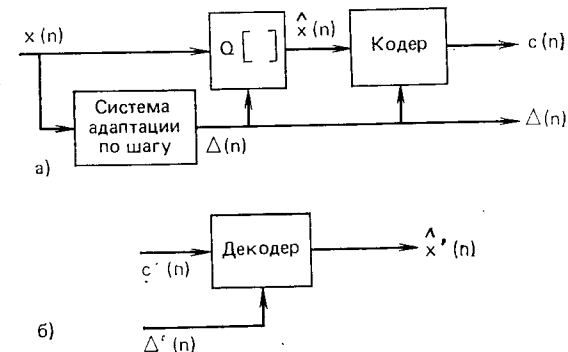


Рис. 5.23. Общая схема адаптивного по входу квантователя: а) кодер; б) декодер

легко можно обобщить на случай неравномерного квантования. Шаг квантования, используемый при квантовании $x(n)$ (рис. 5.23а), должен быть известен на приемной стороне (рис. 5.23б). Таким образом, отсчет описывается кодовым словом и шагом квантования. Если $c'(n) = c(n)$ и $\Delta'(n) = \Delta(n)$, то $\hat{x}'(n) = \hat{x}(n)$, однако если $c'(n) \neq c(n)$ или $\Delta'(n) \neq \Delta(n)$, например имеются ошибки при передаче, то $\hat{x}(n) \neq \hat{x}'(n)$. Конкретное влияние ошибок определяется методом адаптации. На рис. 5.24 изображена общая схема квантователя с адаптацией по входу на основе усилителя с

переменным коэффициентом усиления. В этом случае квантованный сигнал описывается совместно кодовым словом и коэффициентом усиления.

Для того чтобы разобраться в работе схемы квантователя с адаптацией по входу, полезно рассмотреть ряд примеров. В боль-

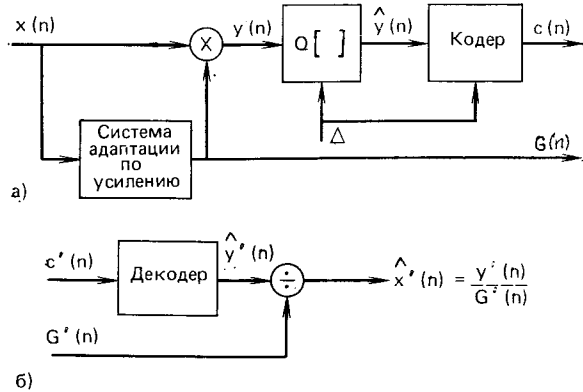


Рис. 5.24. Общий квантователь с адаптацией по входу и с переменным усилением: а) кодер; б) декодер

шинстве систем такого рода используется оценка дисперсии входного сигнала. В этом случае шаг или уровни квантования устанавливаются пропорционально среднему квадратическому отклонению сигнала, а коэффициент усиления — обратно пропорционально.

Общий подход состоит в предложении, что дисперсия пропорциональна кратковременной энергии, которая, как отмечалось ранее, представляет собой сигнал на выходе фильтра нижних частот, на входе которого действует сигнал $x^2(n)$, т. е.

$$\sigma^2(n) = \sum_{m=-\infty}^{\infty} x^2(m) h(n-m), \quad (5.53)$$

где $h(n)$ — импульсная характеристика фильтра нижних частот. [Для стационарного сигнала легко показать (см. задачу 5.7), что математическое ожидание $\sigma^2(n)$ пропорционально σ^2_x .] Например,

$$h(n) = \begin{cases} \alpha^{n-1}, & n \geq 1; \\ 0, & \text{в противном случае,} \end{cases} \quad (5.54)$$

Воспользовавшись (5.54) и (5.53), получаем

$$\sigma^2(n) = \sum_{m=-\infty}^{n-1} x^2(m) \alpha^{n-m-1}. \quad (5.55)$$

Можно показать, что $\sigma^2(n)$ удовлетворяет также уравнению

$$\sigma^2(n) = \alpha \sigma^2(n-1) + x^2(n-1) \quad (5.56)$$

(для устойчивости потребуем, чтобы $0 < \alpha < 1$). Шаг квантования для схемы рис. 5.23а теперь будет равен

$$\Delta(n) = \Delta_0 \sigma(n) \quad (5.57)$$

и переменный параметр адаптации¹

$$G(n) = G_0 / \sigma(n). \quad (5.58)$$

Параметр α определяет протяженность интервала времени, на котором сигнал вносит основной вклад в оценку дисперсии. На рис. 5.25 изображен пример квантователя в системе разностной ИКМ [13]. На рис. 5.25а показана траектория оценки среднего

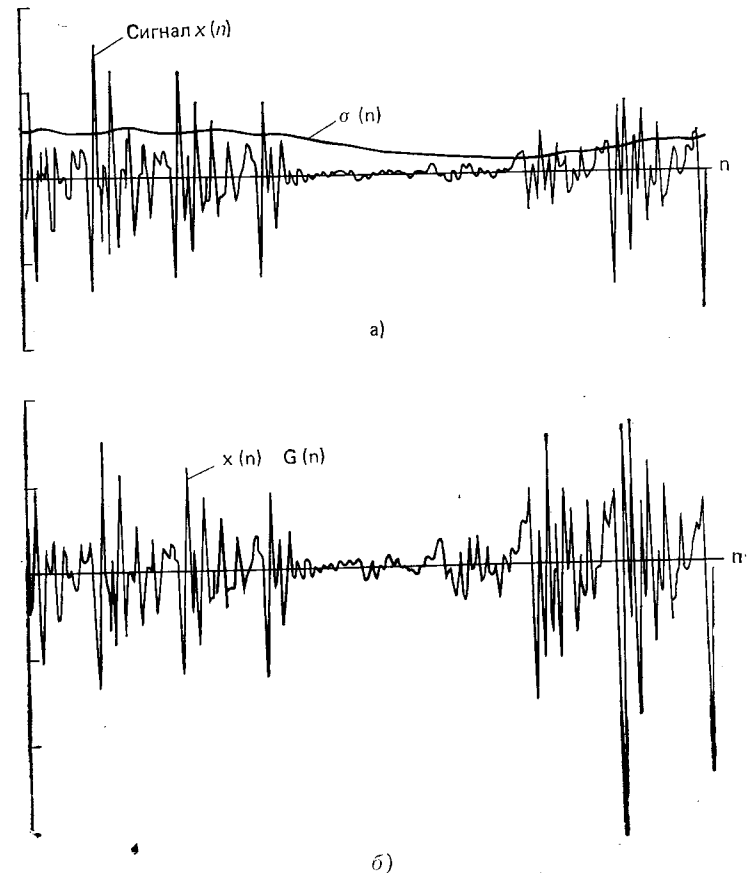


Рис. 5.25. Пример оценки (5.56): а) временная диаграмма сигнала $x(n)$ и $\sigma(n)$ для $\alpha=0,99$; б) произведение переменного коэффициента усиления и сигнала [13]

¹ Предполагается, что константы Δ_0 и G_0 входят в коэффициент усиления фильтра.

квадратического отклонения совместно с входным сигналом $\alpha = 0,99$. На рис. 5.25б показана последовательность $y(n) = x(n)G(n)$. При таком выборе параметра α глубокий провал в амплитуде сигнала не в полной мере компенсируется изменением коэффициента усиления. На рис. 5.26 представлены аналогичные

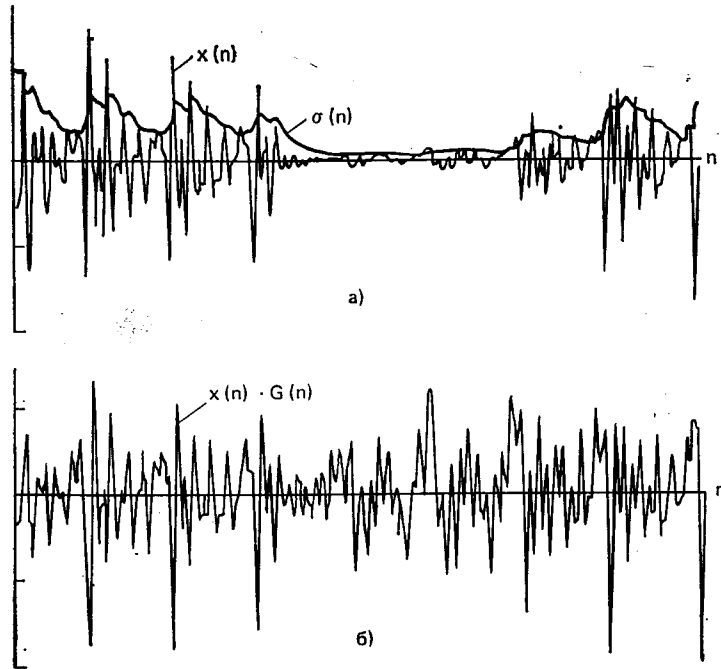


Рис. 5.26. Оценка (5.56):
а) $x(n)$ и $\sigma(n)$ для $\alpha=0,9$; б) произведение $x(n)G(n)$

результаты при $\alpha=0,90$. В этом случае система быстрее реагирует на изменение амплитуды входного сигнала. Таким образом, в этом случае дисперсия $y(n) = G(n)x(n)$ остается постоянной, несмотря на внезапное уменьшение амплитуды входного сигнала. В первом случае при $\alpha=0,99$ постоянная времени (число отсчетов до затухания (e^{-1})) составляет около 100 (или 12,5 мс при частоте дискретизации 8 кГц). Во втором случае при $\alpha=0,90$ постоянная времени равна 9 (или около 1 мс при частоте дискретизации 8 кГц). Таким образом, можно считать, что при $\alpha=0,99$ имеется слововая, а при $\alpha=0,90$ — мгновенная адаптация.

Как следует из рис. 5.25а и 5.26а, оценка среднего квадратического отклонения или обратная ей величина $G(n)$ представляет собой медленно меняющуюся функцию времени по сравнению с исходным сигналом. Частота дискретизации коэффициента усиления (или шага квантования) определяется шириной полосы пропускания фильтра нижних частот. Так, для случая рис. 5.25 и 5.26 частоты, на которых коэффициент усиления фильтра умень-

шается на 3 дБ, равен соответственно 13 и 135 Гц при частоте дискретизации 8 кГц. Важно выбрать наиболее низкую частоту дискретизации, поскольку общая скорость передачи информации складывается из скорости передачи выходного сигнала квантователя и скорости передачи коэффициента усиления. Коэффициент усиления (или шаг квантования) в схемах рис. 5.24 и 5.23 перед передачей следует подвергнуть дискретизации и квантованию.

Перед квантованием следует ограничить диапазон изменения коэффициента усиления или шага квантования. Определим пределы G и Δ интервалом

$$G_{min} \leq G(n) \leq G_{max} \quad (5.59)$$

или

$$\Delta_{min} \leq \Delta(n) \leq \Delta_{max} \quad (5.60)$$

Отношение этих пределов определяет динамический диапазон системы. Для получения примерно постоянного отношения сигнал/шум в диапазоне 40 дБ требуется, чтобы G_{max}/G_{min} или $\Delta_{max}/\Delta_{min} = 100$.

Пример улучшения отношения сигнал/шум при адаптации дан в исследовании, проведенном Ноллом [12]¹. Он рассмотрел алгоритмы с адаптацией по входу и оценкой дисперсии в виде

$$\sigma^2(n) = \frac{1}{M} \sum_{m=n}^{n+M-1} x^2(m). \quad (5.61)$$

Коэффициент усиления (или шаг квантования) передавался через каждые M отсчетов. В данном случае в системе используется буфер объемом M ячеек, и коэффициент усиления, как и шаг квантования, определяется по всем отсчетам, а не только по последнему отсчету, как в предыдущем случае.

В табл. 5.4 приведены результаты сравнения различных трехразрядных квантователей в случае речевого сигнала с известной

Таблица 5.4

Отношение сигнал/шум для адаптивного восьмиуровневого квантователя с адаптацией по входу

Закон	Неадаптивный SNR, дБ	Адаптивный (M=128) SNR, дБ	Адаптивный (M=1024) SNR, дБ
μ-закон ($\mu=100, X_{max}=8\sigma_x$)	9,5	—	—
Гаусса	7,3/6,7*	15,0/14,7	12,1/11,3
Лапласа	9,9/7,4	13,3/13,4	12,8/11,5

* В числителе — для неравномерного, в знаменателе — для равномерного квантователя.

¹ Эти методы исследовались также в [14]. Для обозначения процесса вычисления коэффициента или шага квантования по M отсчетам здесь введен термин «блоковое компандирование».