

РАСЧЕТ ОПТИМАЛЬНЫХ ХАРАКТЕРИСТИК ДЕЛЬТА-КОДЕРА И НИЗКОЧАСТОТНОГО ТРАКТА ДЛЯ ЦИФРОВЫХ СИСТЕМ КОДИРОВАННОЙ ГРОМКОГОВОРЯЩЕЙ СВЯЗИ

Дельта-кодирование является одним из наиболее простых и удобных методов аналого-цифрового преобразования речевых сигналов. Благодаря простым схемотехническим решениям, дельта-кодер может иметь значительно меньшие габариты и меньший ток потребления по сравнению с другими методами аналого-цифрового преобразования. Получаемый код является непрерывным и при передаче по линии связи дельта-кодированного сигнала не требуется передавать дополнительных битов синхронизации, что увеличивает пропускную способность каналов передачи данных.

Недостатком простейшей, линейной дельта-модуляции является то, что оптимальный диапазон входных сигналов, при котором отношение сигнал/шум на выходе декодера максимальный, тон мал и находится в пределах около $\pm 40\%$ [1] от оптимального. Этот недостаток устраняется использованием дельта-модуляции с нелинейным предсказанием или адаптивной дельта-модуляции. В данной статье рассматривается адаптивная дельта-модуляция.

Для расчета основных параметров дельта-модуляции, таких как размер шага квантования, максимальный коэффициент увеличения шага квантования, минимальное отношение сигнал/шум квантования, для выбора тактовой частоты, обеспечивающей эти параметры, потребуется определить оптимальную, с точки разборчивости речи, амплитудно-частотную характеристику (АЧХ) дельта-кодера, минимальное отношение сигнал/шум и динамический диапазон. Чтобы определить указанные исходные значения в применен метод индекса артикуляции. Этот метод был разработан для оценки разборчивости речи по физическим характеристикам тракта передачи. На основании артикуляционных испытаний в 40-е – 50-е годы эмпирически были разработаны специальные графики и таблицы, по которым можно приблизительно оценить разборчивость принимаемой речи, исходя из амплитудно-частотной характеристики тракта передачи и уровня шума. Используемые в статье данные взяты из [2]. В [2] приведена диаграмма для расчета индекса артикуляции, которая показанная на рис. 1.

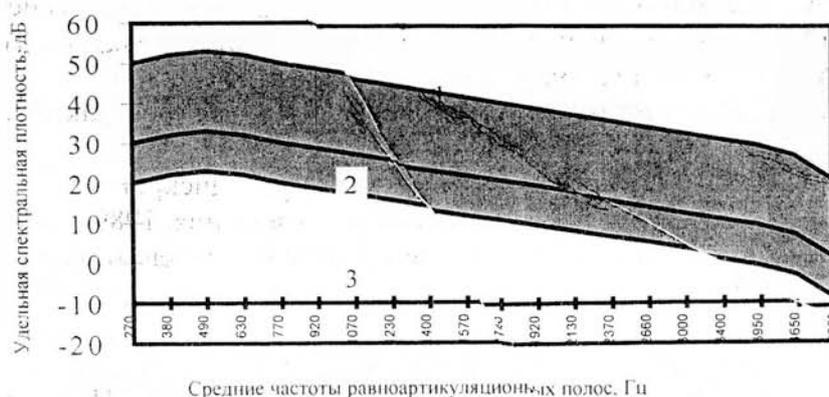


Рисунок 1 – Диаграмма для расчета индекса артикуляции. Общий уровень среднеквадратического давления на расстоянии 1 м от источника речи равен 69 дБ: 1-пиковый уровень речи, 2-средний уровень речи, 3-минимальный уровень речи.

При анализе не будем учитывать искажения сигнала линейным трактом системы передачи данных, трактами усиления входного и выходного аналоговых сигналов, а также предполагаем, что в кодирующем и декодирующем устройствах используется один идеальный интегратор.

На этой диаграмме приведена зависимость уровня спектральной плотности звукового давления на единицу полосы частот. Вся шкала частот разбита на двадцать полос равной артикуляции.

Для обеспечения максимальной разборчивости частотная характеристика тракта должна располагаться выше заштрихованной области, а уровень шумов - ниже этой области. Поэтому оптимальная АЧХ должна как минимум совпадать с верхней границей заштрихованной области (кривая 1). Исходя из этого АЧХ НЧ тракта можно аппроксимировать как АЧХ интегрального типа, приведенную на рис. 2.

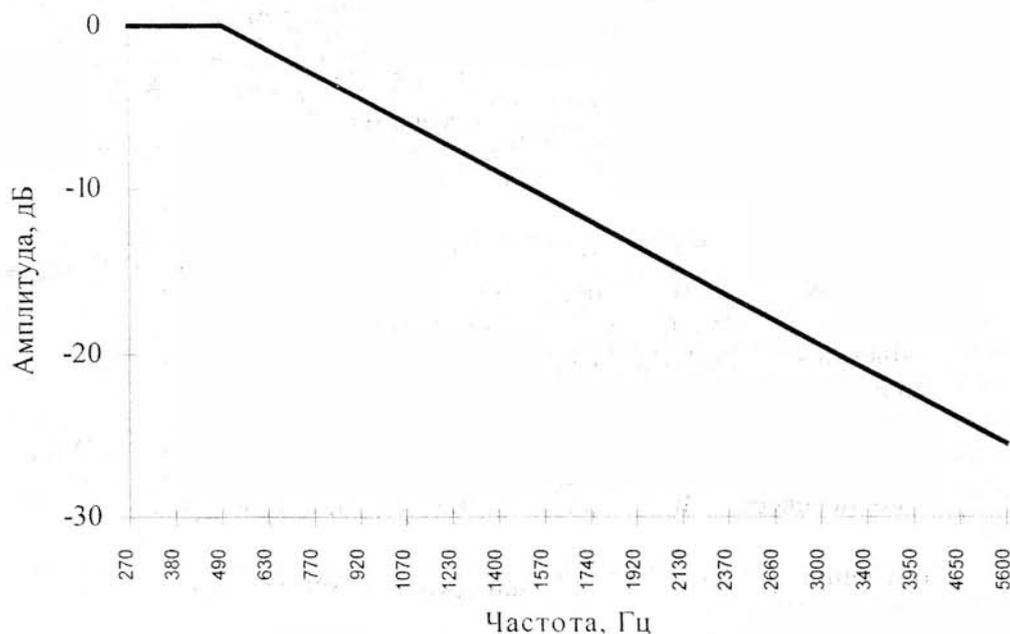


Рисунок 2-АЧХ НЧ тракта.

Частота среза по уровню -3 дБ равна приблизительно 770 Гц. Неравномерность АЧХ (по данным рис. 1) в полосе частот 270-770 Гц относительно частоты 490 Гц должна составлять не более 3 дБ.

Далее частотой среза АЧХ дельта-кодера будем называть частоту, от которой начинается наклон АЧХ, в данном случае 490 Гц.

Учитывая то, что кривая 1, рис. 1, на частотах выше 3950 Гц имеет большую крутизну, чем кривая 1, на рис. 2, при разработке требований к АЧХ тракта мы можем допустить снижение уровня сигнала на частоте 5600 Гц на величину не превышающую значения -5 дБ, относительно заданного.

Необходимо отметить, что дельта-кодер ограничивает по амплитуде только высокочастотные сигналы (перегрузка по крутизне), низкочастотная область полосы частот ограничивается неравномерностью коэффициента усиления усилителей и фильтрами, расположенными до кодера и после декодера.

Исходя из вышесказанного, для оценки АЧХ тракта удобно проводить измерения при уровне входного сигнала, измеренного на частоте 490 Гц, так как при этом уровне на частотах приблизительно до 3000 Гц на искажение спектра влияют только усилители и фильтры. Интегральная частотная характеристика дельта-кодера имеет постоянный

наклон -6 дБ/октаву и если правильно выбрана частота среза дельта-кодера, которая определяется при измерении уровня максимального входного сигнала на этой частоте, то в низкочастотном диапазоне его характеристика будет соответствовать теоретической и для нас будет важнее определить отклонения коэффициентов усиления усилителей и фильтров.

При задании полосы частот НЧ тракта необходимо учесть неидеальность входных фильтров и сделать небольшой запас по частоте, величина которого зависит от выбранных фильтров.

Для удобства округляем значение частоты среза дельта-кодера до величины 500 Гц и получим следующие параметры АЧХ НЧ тракта:

- неравномерность АЧХ в полосе частот 270-3400 Гц при уровне входного сигнала -20 дБ относительно максимального входного сигнала не более 3 дБ;
- неравномерность АЧХ в полосе частот 3400-5600 Гц- не более 10 дБ;
- полоса частот НЧ тракта (частота среза входных фильтров) 250 Гц - 6 кГц.

По данным, приведенным на рис. 1, общий уровень среднеквадратического давления на расстоянии 1 м от источника речи равен 69 дБ.

Как показывают данные, приведенные в [2] при расстоянии 30 см от рта диктора уровень 69 дБ превышает 30 % всех интервалов речи на частоте 500 Гц, а уровень 80 дБ превышает только 1 %.

Так как нас больше интересует использование аппаратуры в ситуациях, когда источник сигнала удален от микрофона на расстояние до нескольких метров, то вполне можно ограничиться минимальным удалением от микрофона на 1 м. В случае приближения объекта к микрофону на расстояние 30 см исказится всего 30 % информации, при этом разборчивость снизится незначительно.

Таким образом, за максимальный уровень звукового давления на входе тракта мы примем значение 69 дБ. Для получения напряжения максимального входного сигнала $U_{\text{пmax}}$ необходимо перевести это значение в паскали и умножить на чувствительность микрофона.

Для получения значения минимального динамического диапазона сигнала воспользуемся диаграммой рис. 1.

Динамический диапазон определим как разность в децибелах между максимальным значением пикового уровня спектральной плотности речи (53 дБ) и минимальным значением минимального уровня спектральной плотности речи (-7 дБ), на рис. 1 он будет равен 60 дБ.

При этом, минимальное отношение сигнал/шум определим как разность между пиковым и минимальными значениями на одной частоте, это значение постоянно во всем частотном диапазоне и равно 30 дБ.

В первую очередь нам необходимо выбрать тактовую частоту такую, чтобы отношение сигнал/шум квантования было больше минимально допустимого (30 дБ).

Для этого нам необходимо определить минимальную мощность шумов квантования, формула которой для интегрального спектра взята из [1]:

$$N_Q = \frac{\pi^2}{6} \left[\frac{(\ln B)^2 + 2.06 \cdot \ln B + 1.17}{B^3} \right] \cdot \left[\frac{\frac{\omega_3}{\omega_m}}{\arctg \frac{\omega_m}{\omega_3}} - \left(\frac{\omega_3}{\omega_m} \right)^2 \right], \quad (1)$$

где B- коэффициент расширения спектра,

ω_3 - угловая частота начала наклона АЧХ,

ω_m - верхняя граница полосы частот.

Коэффициент расширения полосы частот B определяется по формуле [1]:

$$B = \frac{f_s}{2 \cdot f_m} \quad (2)$$

где f_s - тактовая частота дельта-кодера;

f_m - ширина полосы сигнала.

Максимальное отношение сигнал/шум (SNR) для дельта-кодера определяется по формуле [1]:

$$SNR = -10 \lg(N_Q) \quad (3)$$

Так как в формуле (1) используется отношение угловых частот $\frac{\omega_3}{\omega_m}$, которое равно отношению частот $\frac{f_3}{f_m}$, определим заранее это значение для АЧХ. Оно равно $500/6000 = 0,083$.

Подставив известные значения констант и переменных в формулу (1), получим следующую формулу для заданной АЧХ:

$$N_Q = 0,08042 \cdot \left[\frac{(\ln B)^2 + 2,06 \ln B + 1,17}{B^3} \right] \quad (4)$$

Подставив значения в формулы (2), (3) и (4), для разных тактовых частот мы получим значения SNR, приведенные в табл. 1.

Таблица 1

$f_s, \text{кГц}$	B	SNR, дБ
16	1,3	11,8
32	2,6	18,3
64	5,3	24
100	8,3	28,5
128	10,7	33,8
200	16,7	36
256	21,3	38,5
512	42,7	46
1024	85,3	54
2048	170,7	62

На основании данных, приведенных в табл. 1, можно выбрать значение тактовой частоты сигнала f_s . Для нашего случая подходит величина 128 кГц, при тактовой частоте 64 кГц отношение сигнал/шум будет на 20 % ниже минимального, что приведет к потере разборчивости, которую мы будем оценивать далее.

Для этого необходимо рассчитать минимальный шаг квантования и максимальный коэффициент увеличения шага квантования.

Для определения минимального шага квантования воспользуемся формулой [1]:

$$K_{\min} = \frac{\ln 2B}{B} \sqrt{S_1} \cdot \pi \cdot \sqrt{\frac{\frac{\omega_3}{\omega_m}}{\arctg \frac{\omega_m}{\omega_3}} - \left(\frac{\omega_3}{\omega_m}\right)^2}, \quad (5)$$

где $\sqrt{S_1}$ - минимальное среднеквадратическое отклонение сигнала, которое можно приравнять к минимальному уровню входного сигнала ($U_{\text{вх. min}}$).

Подставив в формулу (5) известные параметры АЧХ, получим выражение:

$$K_{\min} = 0,695 \cdot \frac{\ln B}{B} \cdot U_{\text{вх. min}}, \quad (6)$$

где $U_{\text{вх. min}}$ - минимальный уровень сигнала на входе дельта-кодера.

Минимальную крутизну интегратора дельта-кодера можно рассчитать по формуле:

$$\beta_{\min} = k_{\min} \cdot f_s \quad (7)$$

Максимальный коэффициент увеличения шага квантования определяется как отношение максимального входного напряжения к минимальному, которое можно определить, исходя из заданного динамического диапазона. Из заданного значения динамического диапазона D необходимо вычесть максимальное отношение сигнал/шум дельта-кодера и перевести из децибел в отношение напряжений.

$$K_{\max} = 10^{\frac{D - \text{SNR}}{20}} \quad (8)$$

Например, для тактовой частоты 64 кГц (SNR=24 дБ) и динамического диапазона 60 дБ $K_{\max}=63$ раза.

Для определения $U_{\text{вх. min}}$ необходимо сначала задать значение максимального напряжения на входе дельта-кодера для этого можно воспользоваться максимально возможным значением входного напряжения при заданном напряжении питания. Оно определяется по формуле:

$$U_{\max} = \frac{U_n}{2 \cdot \sqrt{2}}, \quad (9)$$

где U_p - напряжение питания дельта-кодера.

Если взять $U_{\max K}$ как верхнюю границу динамического диапазона кодера $U_{\max K}$, то нижней границей динамического диапазона будет $U_{\text{вх. min}}$.

$$U_{\text{вх. min}} = U_{\max K} \cdot 10^{\left(-\frac{D}{20}\right)} \quad (10)$$

Зная $U_{\max K}$ и значение максимального напряжения на выходе микрофона $U_{\text{АП max}}$, мы можем рассчитать коэффициент усиления предварительного усилителя по формуле:

$$K_y = \frac{U_{\max}}{U_{\text{АП max}}} \quad (11)$$

Чтобы оценить разборчивость, необходимо сначала вычислить индекс артикуляции. Для этого совместим на рисунке 1 АЧХ с кривой 1, - это будет верхняя граница перегрузки. Затем, от максимального уровня АЧХ отступим вниз на величину максимального отношения сигнал/шум (SNR) для данной тактовой частоты и нарисуем горизонтальную линию, - это будет нижняя граница шума. После этого вычислим процентное отношение площади заштрихованной фигуры, лежащей между верхней и нижней границами к площади всей заштрихованной фигуры. Эта величина будет называться индексом артикуляции.

Зная величину индекса артикуляции, по графику на рис. 3 [2], можно определить разборчивость речи в процентах.

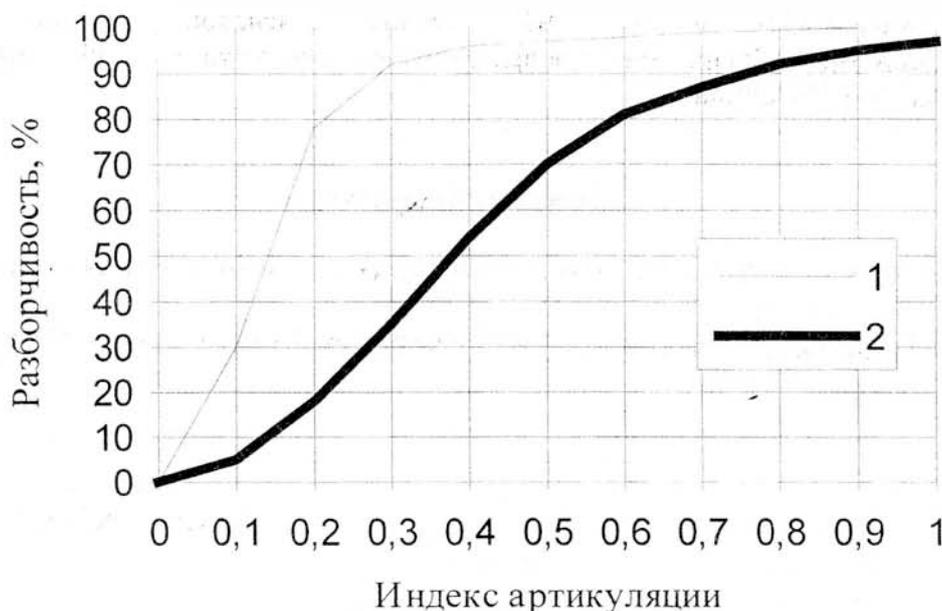


Рисунок 3- Некоторые приближенные экспериментальные соотношения между индексом артикуляции и разборчивостью (Kryter).
1-предложения; 2-бессодержательные слоги (1000 различных слогов).

Для оценки пригодности тракта можно, например, выбрать границу разборчивости 80 %, принятую для телефонных аппаратов.

Результаты расчетов для различных тактовых частот дельта-кодеров приведены в табл.2.

Таблица 2

f _s , кГц	Индекс артикуляции, %	Разборчивость, %	
		предложения	слоги
16	20	76	18
32	40	95	55
64	48	97	67
100	55	98	75
128	65	98	85
1024	98	100	96

Как видно из данных, приведенных в табл. 2, при тактовых частотах дельта-кодера выше 32 кГц качество передачи вполне достаточное для понимания осмысленной речи и при частотах выше 100 кГц качество достаточное для разбора бессодержательных слогов.

Необходимо отметить, что к полученным оценкам разборчивости следует относиться с максимальной осторожностью, так как графики рис. 1 и 3 приблизительны, составлены эмпирически, полученные соотношения сильно зависят от типа материала и тренированности дикторов и слушателей.

Метод индекса артикуляции пригоден только при сравнительно гладких неравномерностях АЧХ и аддитивных шумах со сплошным спектром. При высоких значениях неравномерности АЧХ, шумах с прерывистым спектром и высоким уровнем нелинейных искажений в тракте этот метод неприменим.

Приведенные характеристики и методики могут быть использованы при выборе наиболее оптимальной формы АЧХ, уровней сигналов, тактовой частоты, обеспечивающих минимальные массогабаритные характеристики изделий и минимальную требуемую скорость передачи.

Список литературы:

1. Эбейт (J.E. Abate), Линейная и адаптивная дельта-модуляция, ТИИЭР т55 №3 1967 г.
2. Фланган Дж., Анализ, синтез и восприятие речи, Москва, Связь, 1968 г.

Поступила 28.02.2003г.
После доработки 23.05.2003г.