Вариант 13

ффективное кодирование. Кодирование для ДИБП. Кодовые слова переменной длины. Теорема кодирования. Алгоритм Хаффмена

Реальные источники информации обладают большой избыточностью. Поэтому для ее уменьшения прибегают к эффективному кодированию.
*Эффективное кодировани*е – это процедуры направленные на устранение избыточности.

Эффективное кодирование для ДИБП (дискретный источник информации без памяти):
Пусть ДИБП выдает буквы или символы каждые τ_s секунд. Каждый символ выбирается из конечного алфавита $A \in \{a_k\}, k = 1,2,...,L$ с вероятностью $p(a_k)$. Энтропия такого источника определяется по формуле (2.4) и ограничивается сверху значением, вычисляемым по (4.5), т.е. $H(X) \le \log_2(L)$. Как говорилось выше, знак «=» выполняется, если вероятности символов на выходе источника одинаковы и равны $p = \frac{1}{L}$.

Кодовые слова переменной длины:

Если символы источника не равновероятны, то более эффективно использовать кодовые слова переменной длины.
Пример: *код Морзе (19 век). Символам, возникающим более часто, ставятся в соответствие более короткие кодовые слова, а символам, возникающим менее часто, сопоставляются более длинные кодовые слова.* Такой метод кодирования, который требует знания вероятностей появления символов источника, называется **энтропийным**.

В общем, **префиксное условие** кода требует, чтобы для кодового слова длины K $(b_1...b_M b_{M+1}....b_K)$ не существовало других кодовых слов длины $M < K$ с элементами $(b_1...b_M)$. Это свойство делает кодовые слова однозначно декодируемыми.

Критерий оптимальности однозначно декодируемых кодов переменной длины имеет вид:

$$\overline{K} = \sum_{k=1}^{L} n_k p(a_k) = \min \, , \qquad (4.13)$$

Теорема Шеннона кодирования ДИБП:

Пусть X - ансамбль символов ДИБП с конечной энтропией H(X) и выходными символами из алфавита A={a1 ,....,aL } с вероятностями выхода $p(a_k)$, k = 1,2,.., L .Тогда существует возможность создать код, который удовлетворяет префиксному условию и имеет среднюю длину K , удовлетворяющую неравенству: **H(X) ≤ K < H(X) +1**

Алгоритм Хаффмена:

Рассмотрим метод кодирования Хаффмена, применение которого к любому произвольному ансамблю символов ДИБП обеспечивает получение оптимального по критерию префиксного кода. Алгоритм кодирования Хаффена. Критерий оптимальности кодов Хаффмена – минимум средней длины кодового слова.

Пример:

Критерий оптимальности кодов Хаффмена – минимум средней длины кодового слова (4.13).

Рассмотрим пример. ДИБП выдает символы из алфавита объемом $L = 7$ с вероятностями:
$p(a_1) = 0.2$, $p(a_2) = 0.35$, $p(a_3) = 0.1$, $p(a_4) = 0.3$, $p(a_5) = 0.005$, $p(a_6) = 0.04$, $p(a_7) = 0.005$.

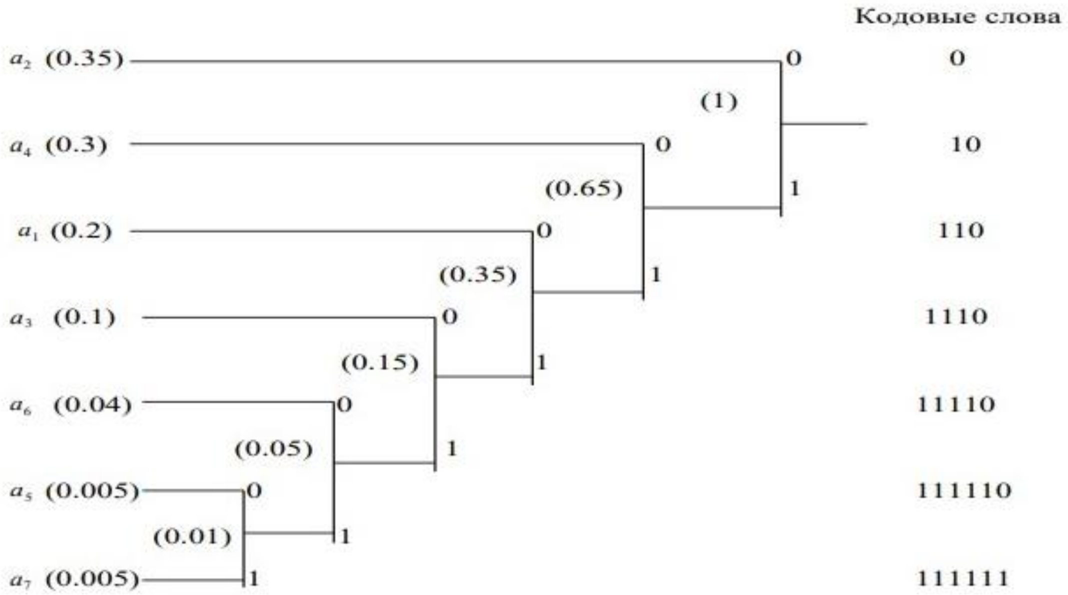
1) Расположить символы источника в порядке убывания (не возрастания) вероятностей.

2) Процесс кодирования начинается с двух наименее вероятных символов a_5, a_7. Эти символы объединяются, причем верхней ветви присваивается «0», нижней «1» или наоборот.

3) Вероятности этих двух ветвей складываются, суммарному узлу присваивается вероятность 0.01.

4) Далее пункты 2), 3) повторяются, пока не исчерпаются символы источника. Вероятность последнего узла равна 1.

Построим кодовое дерево.

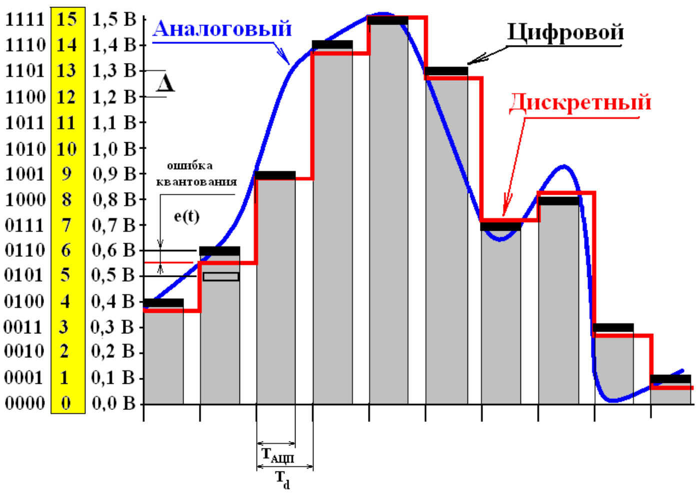


Кодовые слова записываются по кодовому дереву, проходя по нему, справа налево до кодируемого символа. Полученный код не является единственно возможным. Энтропия заданного ДИБП H(X) = 2.11 бит/символ средняя длина кодового слова K = бит/символ

Другие методы:
Альтернативное кодовое дерево кода Хаффмена
Кодовое дерево Хаффмена для посимвольного кодирования
Кодовое дерево Хаффмена для кодирования пар символов

искретизация и восстановление непрерывных сигналов. Теорема Котельникова.

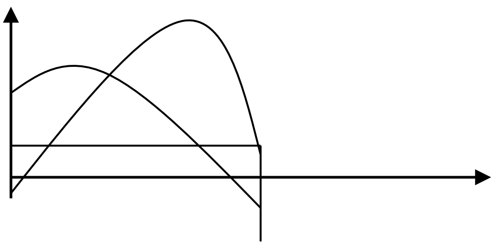
2
.
2



Дискретизация сигнала — замена непрерывного сигнала последовательностью чисел, являющихся представлением этого сигнала по какому-либо базису.

Теорема Котельникова точно справедлива только для сигналов с финитным (конечным) спектром. На рис.3.14 показаны некоторые варианты финитных спектров:

ω) 3



ωв ω

Однако спектры реальных информационных сигналов бесконечны. В этом случае теорема Котельникова справедлива с погрешностью.

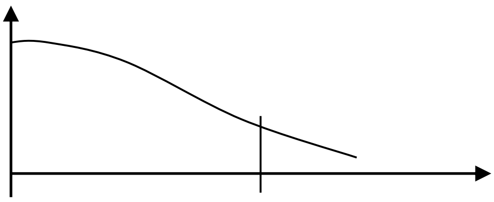
S

x

ω)

ωв ω

Рис.3.15.



Погрешность дискретизации определяется энергией спектральных составляющих сигнала, лежащих за пределами частоты ωв.

$$\overline{\Delta E_\partial^2} = \int_{\omega_в}^{\infty} \left| \dot{S}_x(\omega) \right|^2 d\omega$$

Вторая причина возникновения погрешностей - неидеальность восстанавливающего ФНЧ.

Т.о. погрешность дискретизации и восстановления непрерывного сигнала определяется следующими причинами:

пектры реальных сигналов не финитны.

2) АЧХ реальных ФНЧ неидеальны.

Например, если в качестве ФНЧ использовать RC- фильтр, то восстановленный сигнал на его выходе будет иметь вид:

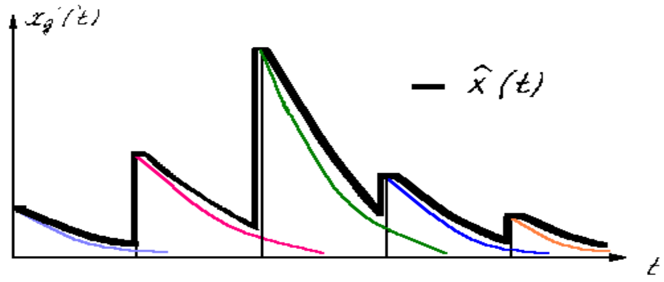


Рис.3.16.

с учетом того, что импульсная реакция RC-фильтра равна: $g_{RC}(t) = \frac{1}{RC} e^{-\frac{t}{RC}}$

Вывод: чем выше $\omega_в$ и чем ближе характеристики ФНЧ к идеальным, тем ближе восстановленный сигнал к исходному.

Теорема Котельникова:

Телекоммуникационные сигналы делятся на непрерывные и дискретные.

Непрерывные сигналы (функции) могут принимать любые, сколь угодно близкие друг к другу значения, в любые моменты времени. Примером непрерывного сигнала является гармоническое колебание.

Дискретные (цифровые) сигналы могут принимать только заранее известные значения, отличающиеся одно от другого на конечную величину, причем изменяться эти значения могут только в определенные моменты времени. Примером дискретного сигнала является (см. рис.2.1) периодическая последовательность прямоугольных импульсов, которая в моменты времени **τ/2** +кТ) принимает значения или 0, или А

Любая непрерывная функция, спектр которой не содержит частот выше $\omega_в$ ***, полностью определяется своими отсчетами, взятыми через интервал времени*** $\Delta t = \pi / \omega_в$***. (Теорема Котельникова)***

адача ↓↓↓↓↓↓↓↓↓

Вариант 13

Задача. Эргодический дискретный по уровню случайный сигнал принимает значения: $x_0 = -2, x_1 = 0, x_2 = 1, x_3 = 2$ с вероятностями $p_0 = 0.3, p_1 = 0.4, p_2 = 0.1, p_3 = 0.2$. Определите математическое (m_x) ожидание, дисперсию (D) и (m_2) среднюю мощность на единичном сопротивлении. Построить одну из возможных его реализаций на фиксированном интервале $T_н$.

n	0	1	2	3
x_n	-2	0	1	2
p_n	0,3	0,4	0,1	0,2

$$m_x = \sum_{0}^{n} x_n p_n = -2\cdot 0{,}3 + 0\cdot 0{,}4 + 1\cdot 0{,}1 + 2\cdot 0{,}2 = -0{,}1$$

$$m_2 = \sum_{0}^{n} x_n^2 \cdot p_n = 4\cdot 0{,}3 + 0 + 1\cdot 0{,}1 + 4\cdot 0{,}2 = 2{,}1$$

$$D = m_2 - m_x^2 = 2{,}1 - 0{,}01 = 2{,}09$$

необязательно