Доклады БГУИР
2007 январь–март № 1 (17)

УКД 621.391(075.8)

ДВУХЭТАПНОЕ ДЕКОДИРОВАНИЕ ТАБЛИЧНЫХ КОДОВ И ИДЕНТИФИКАЦИЯ КРАТНОСТИ ОШИБОК

В.К. КОНОПЕЛЬКО, ФАМ ХАК ХОАН

Белорусский государственный университет информатики и радиоэлектроники
П. Бровки, 6, Минск, 220013, Беларусь

Поступила в редакцию 20 декабря 2006

В статье рассматривается двухэтапный контроль ошибок табличными кодами, основанный на коррекции ошибок и идентификации их кратности внутренним кодом и исправления стираний внешним кодом.

Ключевые слова: табличный, итеративный, код, ошибки, стирания, кодовое расстояние, кратность, кодирование, декодирование, синдром, норма.

Введение

Табличные (итеративные) коды нашли широкое применение в цифровой звукозаписи и телевидении, различных запоминающих системах в силу их высокой корректирующей способности. Это достигается благодаря высокой информационной избыточности и сложности обработки (аппаратной и временной), которая резко возрастает с увеличением длины и кодового расстояния используемых кодов [1–3]. В данной статье исследуется двухэтапное декодирование многократных ошибок табличными кодами на основе идентификации кратности ошибок с последующим переводом ошибок в стирания на первом этапе, их исправления на втором этапе, что позволяет уменьшить сложность вычислений, увеличить быстродействие декодеров.

Структура и декодирование табличных кодов

Табличный код C=C1⊗C2 представляется собой таблицу размерности $n_1 \cdot n_2$, где каждая строка является кодовым словом длины n_1 кода C1 (внутреннего кода), а столбец — кодовым словом длины n_2 кода C2 (внешнего кода) (рис. 1).

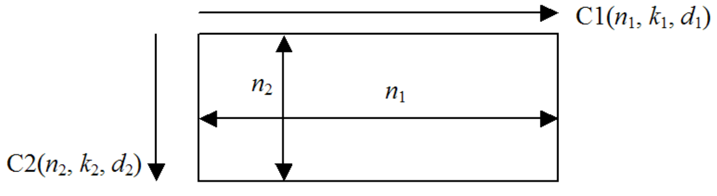

Рис. 1. Структурная схема табличного кода

Известно, что кодовое расстояние табличного кода равно произведению кодовых расстояний кодов C1, C2, т.е. $d_\Sigma = d_1 \cdot d_2$ [1–3]. Поскольку два кода C1, C2 используют одни и те же информационные разряды, имеется возможность двухэтапного декодирования: частичной коррекции ошибок кодом C1 и удаление оставшихся ошибок кодом C2. Рассмотрим случай, когда коды C1, C2 (например, коды Хэмминга) имеют кодовые расстояния d_1=d_2=3, d_Σ=9 и, следовательно, табличный код имеет возможность коррекции любых 4-кратных ошибок. Для карты ошибок (рис. 2,*а*) коррекция может происходить следующим образом: вначале исправляются ошибки кодом С1 (рис. 2,*б*), а затем удаляются ошибки кодом С2 (рис. 2,*в*). Однако при данном алгоритме декодирования часть 4-кратных ошибок может быть не исправлена, например, вида, представленного на рис. 2,*г*.

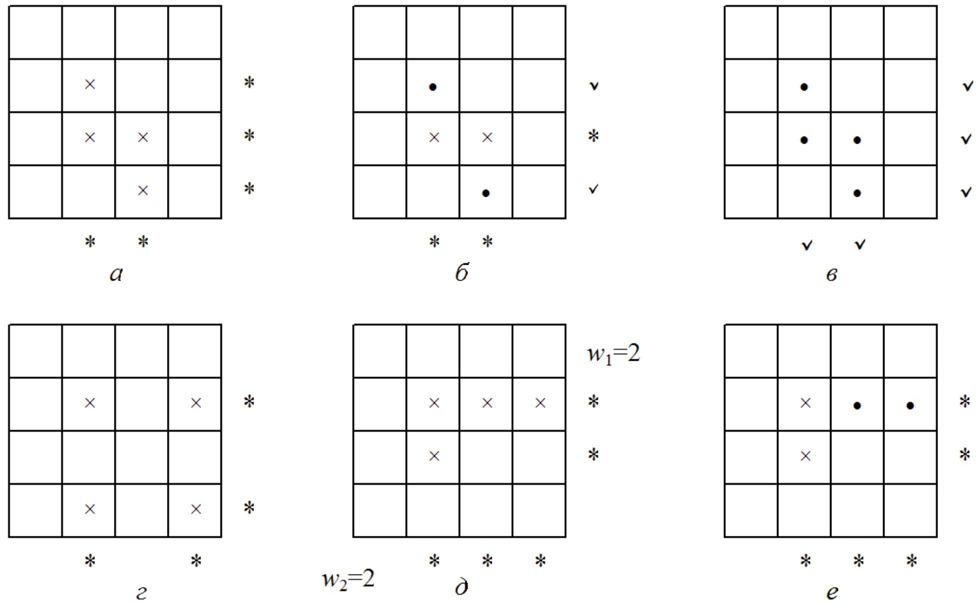

Рис. 2. Карты ошибок и коррекция ошибок табличным кодом

Видно, что на первом этапе декодирования не имеется возможности частичной коррекции ошибок. Это обусловлено тем, что коды С1, С2 не различают (не идентифицируют) по виду синдромов однократных ошибок от двукратных. Кроме того, остается весьма высокой сложность декодирования из-за проблемы селектора (сложности вычисления вектора ошибок по виду синдромов кодов С1, С2) при больших кодовых расстояниях d_1, d_2 и длинах кодов n_1, n_2.

В системах цифрового звука и видеозаписи оптических компакт-дисков, мини- дисков и других систем [2, 4], как правило, декодирование табличных кодов также выполняется в два этапа. На первом этапе декодер кода С1 не только корректирует ошибки небольшой кратности t_1 в каждой строке таблицы (т.е. частично устраняет ошибки), но и помечает (идентифицирует) строки (слова кода С1), у которых число ошибок $t>t_1$. Эти помеченные строки получили название стираний и являются стертыми разрядами столбцов кода С2, декодируемого на втором этапе. Благодаря этому, местоположение стертых разрядов в коде С2 известно, но не известно их истинное состояние. Эта дополнительная информация о местонахождении стертых разрядов в словах кода С2 позволяет в два раза эффективнее использовать корректирующие возможности кодов [1, 2], т.е., например, при использовании кода с d=5 можно исправить четыре стирания t_c=4 вместо двух ошибок t_0=2. Алгоритм двухэтапного декодирования табличных кодов приведен на рис. 3.

На втором этапе декодирование ошибок заменяется на декодирование стираний, которое можно осуществить достаточно просто с использованием переборной процедуры исправления стираний, а именно: стертые символы устанавливаются в состояния "0" и вычисляется синдром S. Если S=0, то принимается, что значения стертых символов установлены правильно. В противном случае, один из стертых символов заменяется на "1" и вновь выполняется вычисление S и т.д., пока соответствующий синдром S_i не будет равен нулю [3]. Очевидно, что при больших кратностях стираний t_c это требует высоких временных затрат как из-за перебора числа состояний стираний, так и задержек на вычисления синдромов.

В [5] предлагается быстродействующий алгоритм стираний, основанный на вычислении всевозможных сумм синдромов одиночных стираний и сравнении их с вычисленным один раз значением синдрома S обрабатываемого слова кода С2. Сложность подобных вычислений суммирования и сравнения r разрядных чисел пропорционально $C_л^c = 2^{t_c+1}$ и не зависит от длины слова n. Это значительно меньше сложности селектора $C_л^o = \sum_{i=0}^{t_0} C_n^i$ при коррекции того же числа ошибок t_c=t_0.

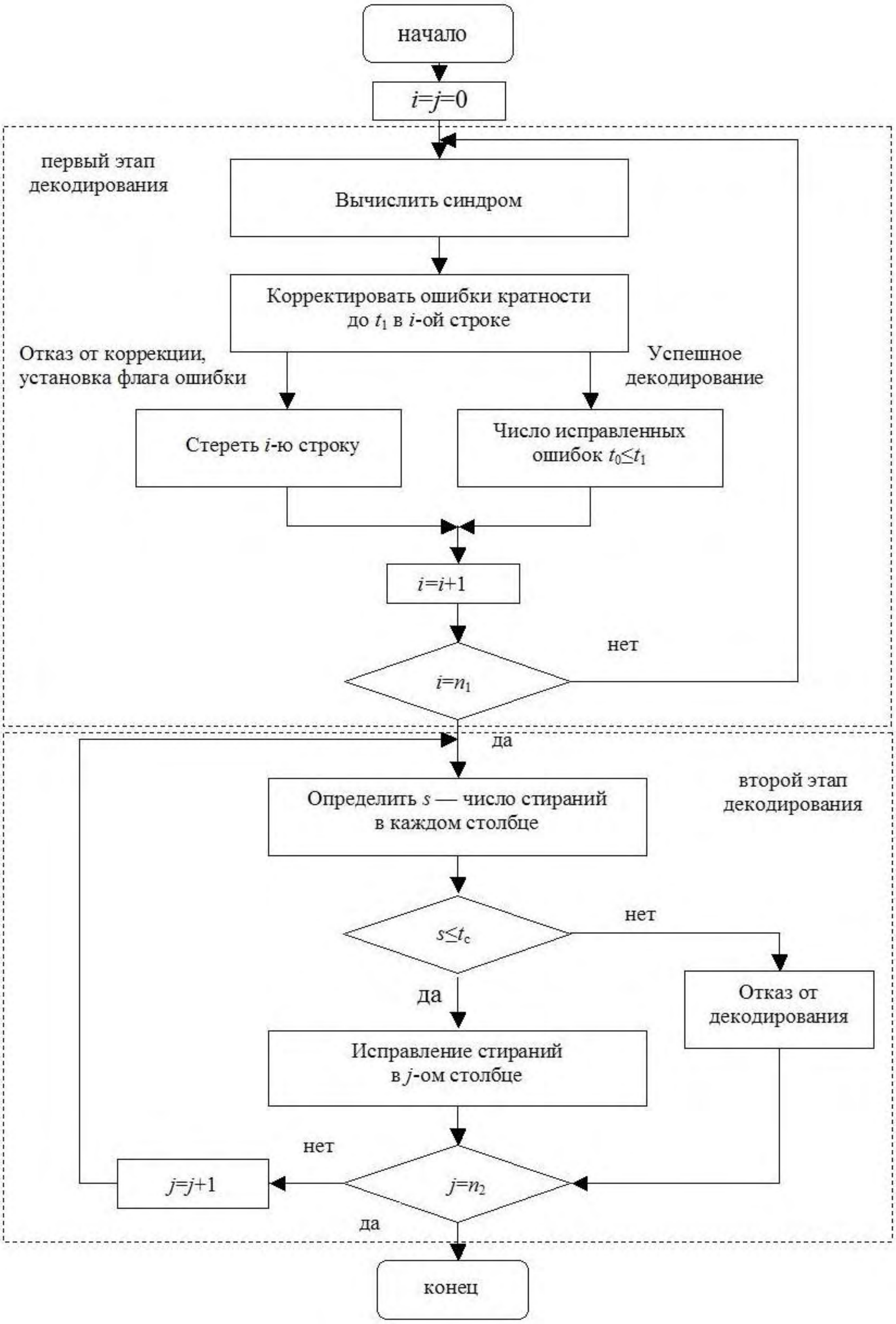

Рис. 3. Алгоритм двухэтапного декодирования табличных кодов

Использование в качестве С1 кода с d_1=4 и С2 кода с d_2=3 позволяет исправить все ошибки, отмеченные на рис. 2,*г*, поскольку данный код С1 позволяет идентифицировать однократные ошибки от двойных. Вместе с тем ошибки, представленные на рис. 2,*д* этот код не идентифицирует, хотя табличный код имеет d_{Σ}=12 (если бы код С1 их идентифицировал как стирания, то с помощью кода С2 они были бы исправлены). Это объясняется тем, что линейные коды позволяют или корректировать ошибки соответствующей кратности (следовательно, имеется возможность идентификации произошедших ошибок, но при этом остается проблема селектора) или же их обнаруживать без идентификации кратности. Поэтому при двухэтапном декодировании корректирующие возможности табличных кодов реализуются не полностью, что определяется прежде всего идентификационными свойствами кода С1, а также выбором алгоритма декодирования.

Идентификация и коррекция ошибок линейными кодами

Вышеприведенное рассмотрение показало, что идентификация кратности произошедших ошибок позволяет увеличить разнообразие видов исправляемых ошибок при двухэтапном декодировании табличных кодов. Рассмотрим возможности кодов по коррекции и идентификации ошибок.

Линейный код с кодовым расстоянием d может одновременно корректировать $t_к$ ошибок и идентифицировать $t_и$ ошибок, для которых

$d \geq t_к+t_и+1$, где $t_и>t_к$. (1)

Это следует из следующего предложения.

Предложение. *Пусть С — линейный* (n, k) *код с кодовым расстоянием* d. *Синдромы векторов, содержащих* $t_к$ *ошибок, отличаются от синдромов векторов, содержащих* $t_и$ *ошибок, если*

$d \geq t_к+t_и+1$. (2)

Доказательство. Пусть e_1 — вектор пространства E_n веса $t_к$, e_2 — вектор пространства $\mathbf{E}_n$ веса $t_и$. Предположим, что синдромы $S(e_1)$, $S(e_2)$ совпадают. Но тогда $S(e_1–e_2)=S(e_1)–S(e_2)=0$, то есть, $e_1–e_2$ — кодовое слово веса не больше $t_к+t_и<d$, чего быть не может. Предложение доказано.

Из предложения следует, что, для того чтобы использовать идентификацию ошибок наряду с коррекцией, кодовое расстояние должно быть не менее 4. Например, удлиненный код Хемминга с d=4 может исправлять одиночные ошибки и одновременно идентифицировать двойные ошибки от одиночных ошибок по виду синдрома. В таблице представлены максимальные идентифицируемые кратности ошибок $t_и$ для различных кодовых расстояний и корректируемых ошибок t_k соответственно. Из таблицы следует, что с уменьшением числа корректируемых ошибок при постоянном кодовом расстоянии d происходит увеличение кратности идентифицируемых кодом ошибок $t_и$.

Кратность корректируемых и идентифицируемых ошибок

d	4	5	6		7		8			9			10			
$t_к$	1	1	1	2	1	2	1	2	3	1	2	3	1	2	3	4
$t_и$	2	3	4	3	5	4	6	5	4	7	6	5	8	7	6	5

Анализ данных таблицы, кроме того, показывает, что, во-первых, для идентификации предпочтительно использование кодов с общей проверкой четности из-за достаточно простой идентификации четных ошибок от нечетных по виду разряда четности; во-вторых, можно увеличить идентифицирующие свойства кода за счет уменьшения возможной кратности корректируемых ошибок. В этом случае число стертых слов (строк) осуществляется со значительно меньшими вычислительными затратами по сравнению с коррекцией этих разрядов как ошибок. Очевидно, что это требует большей информационной избыточности кода С2. При этом допустимое число стертых слов (строк таблицы) $t_с$ определяет кодовое расстояние d_2 кода С2: $d_2=t_c+1$.

Для идентификации кратности ошибок можно воспользоваться недавно разработанной теорией норм синдромов для БЧХ и РС кодов. Для этих кодов по норме синдрома N можно определить, какой класс ошибок соответствующей кратности имеется и, следовательно, идентифицировать кратность ошибок. Например, для БЧХ кода с кодовым расстоянием d=5, задаваемого проверочной матрицей $H=\left[\alpha^z,\alpha^{3z}\right]^T$, синдром $S=\left[\alpha^i,\alpha^j\right]^T$. Классы корректируемых одиночных и двойных ошибок характеризуются нормой синдрома $N=(j-3i) \bmod n$. Для одиночных ошибок норма не зависит от положения ошибок и равна N=0, для двойных — нормы N постоянные, различные и не равны нулю в соответствующем классе ошибок. Если, например, синдром $S\neq 0$ и вычисленная норма $N\neq 0$, то можно считать, что произошли ошибки кратности два или три. При этом достаточно селектировать только нормы N, а не все возможные комбинации синдромов. Кроме того, можно различать между собой часть трехкратных, модульных длины 4, сплошных пакетных ошибок от двукратных [6].

Декодирование с идентификацией числа ошибочных строк и столбцов в коде

Выше отмечалось, что двухэтапное декодирование табличных кодов не позволяет использовать полностью в соответствии с d_Σ корректирующие возможности табличного кода. Их реализация возможна, если отказаться от приоритетной обработки кода С1 на первом этапе декодирования, а обрабатывать первыми слова кода, у которого содержится меньшее число ошибок (т.е. прежде очистить код от ошибок малой кратности).

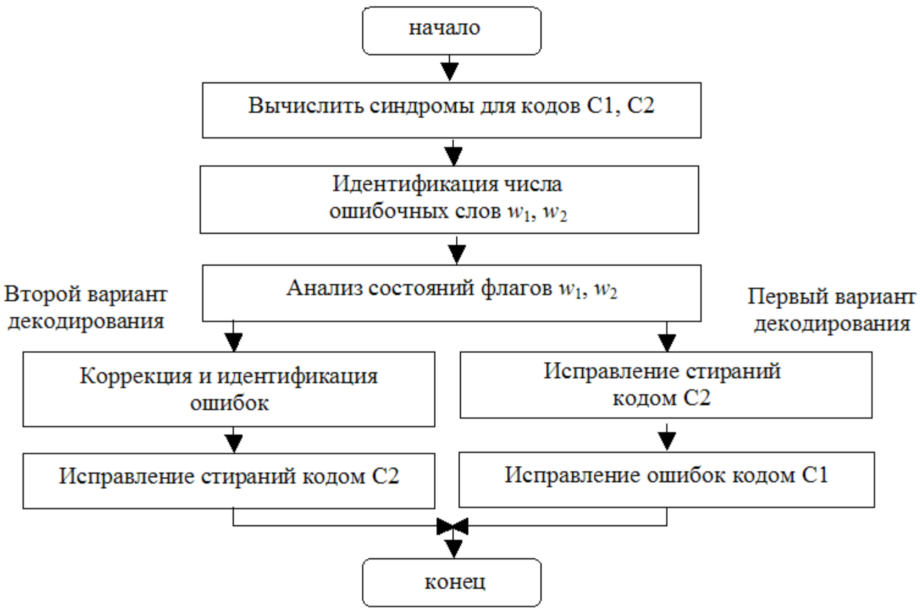

Рис. 4. Алгоритм декодирования с идентификацией весов кодами С1, С2

Анализ показал, что для принятия подобного решения информации об обнаружении и идентификации ошибок, стираниях недостаточно, необходима дополнительная информация о суммарном числе ошибочных слов (весов) w_1 и w_2 в кодах С1 и С2 (это легко вычисляется по $S\neq 0$). Все 4-кратные ошибки исправляются для w_1=2, w_2=2 (рис. 2,*д*) при использовании кодов С1 и С2 с кодовыми расстояниями d_1=5, d_2=2 и d_Σ=10, если первыми будут декодироваться стирания кодом С2, а затем ошибки кодом С1 (рис. 2,*е*). Возможен и второй вариант декодирования: первым обрабатывается слово кода С1 с одиночной ошибкой, а затем оставшиеся стирания корректируются кодом С2 (это реализуемо, поскольку синдромы S^1 одиночных ошибок не совпадают с синдромами S^3 тройных ошибок $S^1\neq S^3$ для двух ошибочных слов кода С1). Алгоритм декодирования с идентификацией числа ошибочных строк и столбцов в коде приведен на рис. 4.

Заключение

Проведенный анализ двухэтапного декодирования табличных кодов показал, что это декодирование эффективно при коррекции ошибок на первом этапе и стираний на втором этапе декодирования. Однако при этом не исправляются определенные виды ошибок, которые код должен корректировать. Показано, что идентификация кратности произошедших ошибок внутренним кодом позволяет дополнительно контролировать некоторые виды ошибок, но требует увеличения информационной избыточности. Рассматривается возможность полного использования корректирующих свойств табличного кода при применении идентификации числа ошибочных строк и столбцов. Приводятся эффективные алгоритмы по коррекции стираний и идентификации кратности ошибок в словах норменными методами.

TWO-STAGE DECODING ITERATED CODES AND IDENTIFICATION MULTIPLICITIY OF ERRORS

V.K. KONOPELKO, PHAM KHAC HOAN

Abstract

In this article two-stage errors control using iterated codes is considered. It is based on errors correction and identification multiplicity of errors by inner code and erasures correction by outer code.

Литература

1. *Конопелько В.К., Лосев В.В.* Надежное хранение информации в полупроводниковых запоминающих устройствах. М., 1986.
2. Теория прикладного кодирования / Под ред. В.К. Конопелько. Минск, 2004. Т. 2.
3. *Кларк Дж., мл., Кейн Дж.* Кодирование с исправлением ошибок в системах цифровой связи. М., 1987.
4. Цифровая звукозапись / Под ред. Дж. Масеса и М. Веркаммена. М., 2004.
5. *Фам Хак Хоан, Конопелько В.К.* // Докл. БГУИР. 2006. Т. 4, №6. С. 19–22.
6. *Конопелько В.К, Липницкий В.А.* Теория норм синдромов и перестановочное декодирование помехоустойчивых кодов. М., 2004.