

Поступила в редколлегия 18 ноября 1994 г.

УДК 681. 32

## АЛГОРИТМ БЫСТРОГО СЖАТИЯ ДВОИЧНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ

Соловей В. А., Кулик И. А., Арбузов В. В.

Методы и средства сжатия стали неотъемлемой частью систем хранения и передачи информации. В настоящее время они имеют особое значение для передачи и хранения видеoinформации, например, в системах телевидения, машинного зрения и факсимильной связи.

Среди основных параметров конкретного сжимающего алгоритма таких, как коэффициент сжатия, помехоустойчивость, схемотехническая реализация, следует выделить время сжатия информации. Его уменьшение необходимо в случае передачи сжимаемой информации в реальном масштабе времени, например, в системах, где отображение данных происходит с частотой порядка 50 Гц. Существующие методы для решения этой задачи хотя и дают хороший коэффициент сжатия, однако они отличаются довольно сложной технической реализацией [1].

В данной работе рассматривается метод сжатия, пригодный, в частности, для телевизионных информационных систем и отличающийся простотой реализации и малым временем преобразования.

Предлагаемый метод основан на возможности представления кодируемой двоичной информации в следующих двух видах: векторном и координатном.

При векторном кодировании двоичная  $n$ -разрядная последовательность может содержать 1, 2, ...,  $n$  логических единиц, а общее число формируемых векторов при этом равно  $2^n$ . Очевидно, что число бит информации, требуемое для передачи и хранения одного вектора, в этом случае

$$I_b = \log_2 2^n = n. \quad (1)$$

При координатном кодировании двоичный вектор преобразуется в последовательность, состоящую из  $k$  координат нулей или единиц. Для этого каждому разряду вектора при счете слева направо присваивают двоичный номер, начиная с нулевого, и затем передают или хранят координаты единичных или нулевых разрядов. Например, двоичному вектору **10001010** соответствует последовательность координат единичных разрядов **000.100.110**. Число бит информации, требуемое для получения последовательности из  $k$  координат,

$$I_k = k \log_2 n. \quad (2)$$

При этом следует учесть, что для уменьшения избыточности координатного метода при числе единиц  $k > n/2$  необходимо кодирование единиц заменить кодированием  $(n - k)$  нулей и наоборот.

Так как число возможных векторов с  $k$  логическими единицами (нулями) равно сочетанию  $C_n^k$ , то реальное количество информации, содержащееся в двоичном векторе с известным числом  $k$ ,

$$I_p = \log_2 C_n^k. \quad (3)$$

Величина избыточности для векторного  $D_v$  и координатного  $D_k$  методов соответственно равна :

$$D_b = 1 - \frac{I_p}{I_n} = 1 - \frac{\log_2 C_n^k}{n}; \quad (4)$$

$$D_k = 1 - \frac{I_p}{I_n} = 1 - \frac{\log_2 C_n^k}{k \log_2 n}. \quad (5)$$

Из сравнительного анализа кодирующих методов следует, что подход к рациональному кодированию сообщения в нашем случае включает процедуру выбора метода кодирования, критерием которого является меньшее значение избыточности. Это означает, что для исходного  $n$ -разрядного вектора с  $k$  логическими единицами необходимо выполнить условие:

$$\frac{n}{k \log_2 n} > 1, \quad (6)$$

а при  $k > n/2$

$$\frac{n}{(n-k) \log_2 n} > 1. \quad (7)$$

Определим область значений  $k$ , при которых выполняются неравенства (6, 7):

$$k < \frac{n}{\log_2 n}; \quad (8)$$

$$k > n - \frac{n}{\log_2 n}. \quad (9)$$

Отсюда общее условие рационального кодирования относительно числа  $k$  имеет вид

$$\frac{n}{\log_2 n} > k > n - \frac{n}{\log_2 n}. \quad (10)$$

Следовательно, если число  $k$  нулей или единиц двоичного вектора удовлетворяет условию (10), то производится его преобразование в последовательность координат. Коэффициент сжатия при этом

$$K'_{сж} = \frac{n}{k \log_2 n}. \quad (11)$$

В противном случае изображение исходного двоичного вектора остается без изменений.

Вышеприведенные выражения справедливы при рассмотрении отдельно взятого двоичного вектора. В случае сжатия массива векторов для их однозначного декодирования необходимо формирование служебных слов, указывающих на количество логических нулей или единиц в векторе. На основании служебных слов декодирующее устройство приемника определяет метод кодирования двоичной комбинации и ее ожидаемую длину. Так как число  $k$  лежит в пределах от 0 до  $n$ , то разрядность служебного слова

$$z = \log_2 n. \quad (12)$$

Использование служебного слова с разрядностью  $z$  приводит к необходимости передачи или хранения

$$I_b = n + \log_2 n. \quad (13)$$

бит информации для векторного метода и

$$I_k = k \log_2 n + \log_2 n = (k+1) \log_2 n. \quad (14)$$

$$\frac{n}{(k+1)\log_2 n} > 1; \quad (15)$$

$$\frac{n}{(n-k+1)\log_2 n} > 1. \quad (16)$$

Преобразуя выражения ( 15, 16 ) относительно числа  $k$ , находим

$$k < \frac{n}{\log_2 n} - 1; \quad (17)$$

$$k > n - \frac{n}{\log_2 n} + 1. \quad (18)$$

Тогда общее условие рационального кодирования

$$\frac{n}{\log_2 n} - 1 > k > n - \frac{n}{\log_2 n} + 1. \quad (19)$$

Левая и правая части неравенства ( 19 ) являются соответственно нижним и верхним условиями сжатия двоичных векторов. Коэффициент сжатия двоичного вектора с учетом служебного слова

$$K''_{сж} = \frac{n}{(k+1)\log_2 n}. \quad (20)$$

Из вышеприведенных рассуждений следует алгоритм рационального кодирования двоичных векторов:

1. Определяются нижнее и верхнее условия сжатия для  $n$  - разрядных двоичных векторов в соответствии с выражением (19).

2. Подсчитывается число  $k$  единиц в векторе.

3. Сравнивается  $k$  с вычисленными верхним и нижним условиями сжатия. Если  $k$  находится между ними или равно одному из них, то двоичный вектор остается без изменений. Если число  $k$  больше верхнего или меньше нижнего условия сжатия, то двоичный вектор представляется координатами логических нулей или единиц соответственно.

4. Формируется служебное слово разрядности  $\lceil \log_2 n \rceil$ , указывающее число  $k$  единиц в исходном двоичном векторе, и им дополняется передаваемая или хранимая двоичная комбинация.

5. Переход к пункту 2.

Таким образом, рассмотренные особенности алгоритма рационального кодирования двоичных векторов позволяют сократить время преобразования информации, что при достижении приемлемых коэффициентов сжатия означает возможность эффективного использования предлагаемого метода для систем передачи и хранения данных, особенно в телевизионных системах.

## SUMMARY

*The method of data compression using the features of the vector and coordinate coding methods is proposed in this article. It is based on reciprocal transitions from one coding method to another. The supposed method takes little time when coding and decoding binary messages and has good effect when compressing. It is effective for compression of vlewdata.*

## СПИСОК ЛИТЕРАТУРЫ

1. Милт Ленард. Рассмотрение возможных способов сжатия изображений  
// Электроника. - 1993, N 3-4, стр. 10-14.

*Поступила в редколлегию 18 ноября 1994 г.*