

УДК 621.391

Ю. Н. АЛЕКСАНДРОВ, канд. техн. наук, *К. Е. ГЛАЗИН*, *Э. Н. КОРОЛЬ*,
Ю. Ю. МИЛОНОВ, *А. В. ТОВАРНИЦКИЙ*, канд. техн. наук

**ИССЛЕДОВАНИЕ СЛОЖНОСТИ РЕАЛИЗАЦИИ МЕТОДА
ПРЕДСТАВЛЕНИЯ ЦЕЛОЧИСЛЕННЫХ И ВЕЩЕСТВЕННЫХ МАССИВОВ
ДАнных В ПОЛИАДИЧЕСКОЙ СИСТЕМЕ СЧИСЛЕНИЯ (ПСС)**

В любой системе сбора, хранения и переработки информации производятся многократные преобразования исходных массивов данных, что обуславливает необходимость развития методов преобразования. Известно [1], что для оценки эффективности реализации любого преобразования (сжатия) $\varphi_i : X \rightarrow Y_i$ (1) исходного массива данных

из одной системы представления в другую Y_i и обратно [2] объемом соответственно по $s(x)$ и $s(Y_i)$ пользуются следующими соотношениями:

$$\begin{aligned} \eta_i &= (\varphi_i, \varphi_i^{-1}) \leq \eta_0; \\ \kappa_i &= s(X)/s(Y_i) \geq 1; \\ \varepsilon_i &= (\varphi_i, \varphi_i^{-1}) \leq \varepsilon_0, \end{aligned} \quad (2)$$

где η_i — коэффициент сложности преобразований прямого и обратного φ_i^{-1} массивов данных; η_0 — предельное значение коэффициента сложности преобразования, при котором имеет смысл эффективность преобразований φ_i и φ_i^{-1} цифровых массивов данных; κ_i — коэффициент сжатия исходного массива данных x ; ε_i — погрешность преобразований φ_i , φ_i^{-1} массива данных; ε_0 — предельное значение погрешности преобразований φ_i и φ_i^{-1} , при котором имеет смысл эффективность преобразования цифровых массивов данных.

Возможности применения ПСС для сжатия целочисленных и вещественных массивов данных и детальная оценка κ_i приведены в работе [3].

Под сложностью преобразования обычно понимают [4] количество элементарных вычислительных операций, затрачиваемых на прямое φ_i и обратное φ_i^{-1} преобразования массивов данных, дополнительные затраты памяти ЭВМ, необходимые для хранения команд программы, применяющей преобразования.

С учетом преобразований φ_i и φ_i^{-1} коэффициент сложности преобразования η_i можно представить с помощью соотношения

$$\eta_i(\varphi_i, \varphi_i^{-1}) = \eta_i(\varphi_i) + \eta(\varphi_i^{-1}). \quad (3)$$

Очевидно, что коэффициент сложности можно снизить как на этапе кодирования, так и на этапе декодирования. Ниже рассматривается возможность снижения коэффициента сложности на первом этапе.

Всякий целочисленный массив, представленный в виде матрицы $A = (a_{ij})$, $i = \overline{1, m}$; $j = \overline{1, n}$, с помощью преобразования

$$N_j = \sum_{i=1}^m a_{ij} \rho_i, \quad (4)$$

где

$$\rho_i = \prod_{k=i+1}^m \lambda_k, \quad (5)$$

$$\lambda_i = 1 + \max_j (a_{ij}), \quad (6)$$

можно заменить двумя векторами $N = \{N_j\}$ и $\Lambda = \{\lambda_i\}$, по компонентам которых возможно обратное преобразование φ^{-1} :

$$a_{ij} = [N_j/\rho_i] - [N_j/\lambda_i \rho_i] \lambda_i, \quad (7)$$

позволяющее однозначно восстанавливать любой элемент с погрешностью $\varepsilon = 0$ [3].

Для получения количественной оценки рассматриваемых преобразований φ_i и φ_i^{-1} запишем исходный цифровой массив вещественных чисел [6] $B = (b_{ij})$, где $i = 1, m; j = 1, n$ в виде элементов матрицы вектора

$$h_j = \sum_{i=1}^m (b_{ij} - \min_i B) [\prod_{k=i+1}^m (1 + (\max_i B - \min_i B)/d_k)] / d_i, \quad (8)$$

где d_i — шаг дискретизации элементов i -й строки, определяющей минимально возможное приращение числовой последовательности для массива B . Затем элементы h_j с помощью известных соотношений [5] представим в виде

$$\{d_i\} = D; \quad (9)$$

$$\min_i B = \{c_i\} = C; \quad (10)$$

$$(1 + (\max_i B - \min_i B)/d_i) = \{\lambda_i\} = \Lambda. \quad (11)$$

Обратное преобразование φ_i^{-1} выглядит для b_{ij} как

$$b_{ij} = c_i + d_i (\{h_j / \prod_{k=i+1}^m \lambda_k\} - \{h_j / \prod_{i=1}^m \lambda_i\} \lambda_i). \quad (12)$$

Для получения более простого алгоритма сжатия исходных целочисленных и вещественных массивов данных, отвечающих критерию (2) снижения сложности преобразования, целесообразнее получение векторного массива N в виде совокупности элементов вектора $\{N_j\}$. Так, в частности, выражение для сжатия целочисленного массива приобретает вид

$$\{N_j\} = \{\rho_{k+1}\} \times A, \quad (13)$$

если при этом число элементов строки умножаемой матрицы $\{\rho_k\}$ не равно числу элементов столбца матрицы A , то для реализации правила умножения матриц необходимо в качестве недостающего элемента строки дописать «1» (единицу).

На примере матрицы A поясним предложенный алгоритм сжатия данных и осуществим оценку затрат на его реализацию с помощью соотношений (5) — (7), (13):

$$A = \begin{pmatrix} 1 & 0 & 2 & 3 & 0 & 1 \\ 2 & 0 & 0 & 2 & 1 & 1 \\ 0 & 3 & 0 & 4 & 0 & 1 \\ 1 & 2 & 1 & 1 & 1 & 2 \end{pmatrix}.$$

Согласно (6) вектор $\Lambda = \{4, 3, 5, 2\}$ и (13) вектор $\{N_j\} = \{17, 5, 37, 13, 21, 9\}$. Для представления векторов Λ и $\{N_j\}$ требуется 14 знакомест вместо $4 \times 6 = 24$ выделенных знакомест для элементов матрицы A . Восстановление любого элемента $Q_{ij} \in A$ осуществляется с помощью соотношения (7). Так, например, согласно (7) элемент $Q_{2,4}$ равен

$$Q_{2,4} = [13/3 \cdot 4] - [13/3 \cdot 3 \cdot 4] \cdot 3 = 1.$$

Для случая вещественного массива B поставленная задача сводится к упрощению сложности нахождения h_j путем приведения вещественного массива B к целочисленному массиву B^* :

$$B = f(\min_i B, d_i) = f_{D, C}(B^*), \quad (14)$$

где $i = \overline{1, m}$; $j = \overline{1, n}$ или

$$B^* = \{b_{ij}^*\} = \{\{b_{ij}\} - \min_j B\} / d_i. \quad (15)$$

Далее, с массивом B^* , полученным в результате преобразования (15), выполняют сжатие в соответствии с приведенными выше соотношениями (5), (6), (13), применяемыми для сжатия целочисленных массивов; восстановление же производится с помощью соотношения (12).

Проведенные исследования сложности реализации метода представления целочисленных и вещественных массивов данных в ПСС позволяют сделать следующие выводы.

При сжатии данных возможно уменьшение сложности алгоритма нахождения вектора $\{N_j\}$ в результате применения правил умножения матриц.

Определен более простой, а значит, снижающий сложность реализации способ сжатия вещественных массивов данных (элементов вектора Λ) за счет приведения его к целочисленному виду с помощью соотношений (5), (6), (12), (13), (15).

Для получения количественных оценок (размеров массива) реализации необходимы дальнейшие исследования.

Список литературы: 1. Новик А. А. Эффективное кодирование. М., 1965. 2. Об эффективном представлении геологических данных нумерическими функциями / В. А. Амелькин, Ю. А. Воронин // Геологическая информация и математическая геология. М., 1976. С 143—152. 3. Александров Ю. А., Король Э. Н., Милонов Ю. Ю. Исследование эффективности одного метода сжатия массивов целочисленных данных // Радиотехника. 1989. Вып. 92. 4. Витушкин А. Г. Оценка сложности задачи табулирования. М., 1969. 5. Сжатие массивов вещественных данных на основе применения метода полиадических чисел / Ю. Ю. Милонов, А. В. Товарницкий, Э. Н. Король и др. // АСУ и приборы автоматики. 1989. Вып. 96. 6. Акушский И. Я., Заболоцкий В. Н. О комбинаторном подходе и идее сжатия информации // Цифровая вычисл. техника и программирование. 1971. Вып. 6. С. 5—17.

Поступила в редколлегию 31.08.89