УДК 629.391

В.В. БАРАННИК

Харьковский военный университет, Украина

МЕТОД БИНОМИНАЛЬНО – ПОЛИАДИЧЕСКОЙ НУМЕРАЦИИ ДВОИЧНЫХ ДАННЫХ

Излагается метод нумерации двоичных данных на основе формирования кода-номера двухпризнаковому биноминальному числу в полиадическом пространстве.

двоичные данные, биноминальное число, полиадическое пространство, сжатие информации, нумерация, структура

Введение

В соответствии с Национальными космическими программа и Национальной программой информатизации Украины основными заданиями являются решение общемировых и внутриотраслевых задач в реальном времени. При этом возникает ***проблема,*** состоящая в невозможности передать по существующим каналам связи необходимые объемы достоверной информации в заданных временных интервалах. Одно из направлений решения заключается в организации компактного представления данных.

Анализ известных методов компактного представления с контролируемой потерей качества выявил то, что наибольшие коэффициенты сжатия достигаются при устранении структурных видов избыточности [1, 2]. Однако, эти методы не обеспечивают требуемой степени сжатия, достаточной для передачи данных по каналам связи в реальном времени. В большинстве структурных методах ключевым звеном является принцип нумерации. Он включает в себя систему правил, по которой обрабатываемой последовательности ставится в соответствие код-номер. Система правил учитывает особенности, выявленные на последовательности данных по заданным структурным характеристикам. Значит система нумерующих правил (нумератор) отражает суть выявленных особенностей, на основе которых устраняется избыточность и достигается сжатие. Следовательно, недостаточные степени сжатия объясняются в отсутствии нумерующих систем, учитывающих информативные структурные характеристики. Поэтому одним из вариантов разработки новых методов сжатия на основе уменьшения структурной избыточности является создание принципиально новых нумераторов, на основе обобщения достоинств и устранения недостатков свойственных существующим нумераторам.

Цель статьи. Разработать нумерацию двоичных данных на основе обобщения достоинств биноминального и полиадического типов нумераторов.

Разработка биноминально – полиадической нумерации

Рассмотрим множества $\Psi(\mathbf{m},\Lambda)$ и $\Psi(\mathbf{m},\Theta)$, которые образуются соответственно полиадическими $\Lambda = \{\lambda_i\}_{i=\overline{1,m}}$ и биноминальными ограничениями Θ для двоичной последовательности длиной $\mathbf{m}$ элементов [3 - 6]. Биноминальными ограничениями называются такие признаковые (структурные) ограничения, которые разбивают исходное множество натуральных чисел на несколько не пересекающихся классов множеств [5, 6]. Тогда под биноминальным (Б) или полиадическим (П) пространством понимается множество пронумерованных чисел, удовлетворяющих соответственно биноминальным или полиадическим ограничениям. Мощности таких пронумерованных множеств $\Psi(\mathbf{m},\Lambda)$ и $\Psi(\mathbf{m},\Theta)$ соответственно равны $\mathbf{V}(\mathbf{m},\Lambda)$ и $\mathbf{V}(\mathbf{m},\Theta)$. В этом случае при формировании кода-номера для двоичной последовательности $\mathbf{A}=\{\mathbf{a_i}\}_{\mathbf{i}=\overline{\mathbf{1,m}}}$ его значения для полиадических или биноминальных ограничений будет находится соответственно в диапазонах

$$\mathbf{0} \le \mathbf{N}(\mathbf{m},\Lambda) \le \mathbf{V}(\mathbf{m},\Lambda) = \left(\prod_{\mathbf{i}=1}^{\mathbf{m}} \lambda_{\mathbf{i}}\right) - \mathbf{1}\,; \qquad (1)$$

$$\mathbf{0} \le \mathbf{N}(\mathbf{m},\Theta) \le \mathbf{V}(\mathbf{m},\Theta)\,, \qquad (2)$$

где $\mathbf{N}(\mathbf{m},\Lambda)$, $\mathbf{N}(\mathbf{m},\Theta)$ - коды-номера для двоичной последовательности $\mathbf{A}$, на которой выявлены ограничения соответственно полиадического и биноминального характера.

Выберем в качестве биноминального признака – структурную характеристику ϑ число серий единиц. Данная характеристика является наиболее информативной среди известных [3, 5] (обеспечивает наименьшее значение $\mathbf{V}(\mathbf{m},\Theta)$). Тогда выражение (2) примет вид [4, 5]:

$$\mathbf{0} \le \mathbf{N}(\mathbf{m},\Theta) \le \frac{(\mathbf{m}+\mathbf{1})!}{(2\vartheta)!(\mathbf{m}+\mathbf{1}-2\vartheta)!} - \mathbf{1}\,. \qquad (3)$$

Для уменьшения значений $\mathbf{V}(\mathbf{m},\Lambda)$ и $\mathbf{V}(\mathbf{m},\Theta)$, а, следовательно, и снижения значений кодов $\mathbf{N}(\mathbf{m},\Lambda)$ и $\mathbf{N}(\mathbf{m},\vartheta)$ предлагается построить множество $\Psi(\mathbf{m},\Lambda,\vartheta)$ двоичных последовательностей, удовлетворяющих одновременно ограничениям Λ и ϑ. В этом случае последовательность $\mathbf{A}$ будет принадлежать множеству $\Psi(\mathbf{m},\Lambda,\vartheta)$ тогда, когда через позицию с $\lambda_{\mathbf{i}}=\mathbf{1}$ не будет проходить серия единиц, т.е. полиадические ограничения трактуются как запрет появления единиц на определенной позиции. Значит с одной стороны на расположение серий единичных элементов накладываются дополнительные условия, задаваемые полиадическими ограничениями $\mathbf{0} \le \mathbf{a_i} \le \lambda_{\mathbf{i}} - \mathbf{1}$. С другой стороны, не все двоичные последовательности, удовлетворяющие полиадическим ограничениям, будут принадлежать множеству $\Psi(\mathbf{m},\Lambda,\vartheta)$. Ведь там где разрешена единица с точки зрения полиадических ограничений может стоять разрыв (значение равно $\mathbf{0}$) между двумя соседними сериями единиц, т.е. единица будет запрещена с точки зрения биноминальных ограничений. Отсюда следует выполнение системы неравенств:

$$\mathbf{V}(\mathbf{m},\Lambda,\vartheta) \le \mathbf{V}(\mathbf{m},\Lambda)\,; \qquad (4)$$

$$\mathbf{V}(\mathbf{m},\Lambda,\vartheta) \le \mathbf{V}(\mathbf{m},\vartheta)\,, \qquad (5)$$

где $\mathbf{V}(\mathbf{m},\Lambda,\vartheta)$ - количество двоичных последовательностей, принадлежащих множеству $\Psi(\mathbf{m},\Lambda,\vartheta)$.

Кодирование двоичных последовательностей, принадлежащих множеству $\Psi(\mathbf{m},\Lambda,\vartheta)$, организуется на основе системы правил, ставящих во взаимооднозначное соответствие последовательность $\mathbf{A}\in\Psi(\mathbf{m},\Lambda,\vartheta)$ и значение кода-номера $\mathbf{N}(\mathbf{m},\Lambda,\vartheta)$.

Для разработки такого кодирования сформулируем определение и докажем теорему о нумерации.

Определение 1. Биноминально – полиадической (БП) нумерацией двоичных данных называется система выражений, формирующих код-номер для биноминального числа в двоичном полиадическом пространстве или формирующая код-номер для двоичного полиадического числа в биноминальном пространстве.

В связи с этим под БП множеством понимается некоторая часть Б и П множества, которая соответствует одновременно биноминальным и полиадическим ограничениям. Следовательно, принадлежащие множеству $\Psi(\mathbf{m},\Lambda,\vartheta)$ двоичные последовательности являются биноминально – полиадическими числами.

Теорема о биноминально – полиадической нумерации. Для любой двоичной последовательности

$A=\{a_i\}_{i=\overline{1,m}}$ с выявленными ограничениями на позиции единиц $\Lambda=\{\lambda_i\}_{i=\overline{1,m}}$ и числом ϑ серий единиц можно сформировать код-номер $N(m,\Lambda,\vartheta)$:

$$N(m,\Lambda,\vartheta)=\sum_{k=1}^{K} N(\Theta^{(k)})=\sum_{k=1}^{K}\prod_{z=1}^{Z} N\left(\vartheta_z^{(k)},\,\Theta^{(k)}\right); \quad (6)$$

$$N\left(\vartheta_z^{(k)},\,\Theta^{(k)}\right)=\sum_{i=1}^{m_z} a_i \times$$
$$\times \sum_{\chi_z^{(k)}=\vartheta_z^{(k)}}^{m_z-\vartheta_z^{(k)}+1} \binom{\chi_z^{(k)}-1}{\vartheta_z^{(k)}-1}\binom{m_z-\chi_z^{(k)}+1}{\vartheta_z^{(k)}}, \quad (7)$$

где $\vartheta_z^{(k)}$ - значение числа серий для z-ой допустимой зоны двоичной последовательности A;

$\chi_z^{(k)}$ - число единиц, размещаемое в $\vartheta_z^{(k)}$ сериях;

$\Theta^{(k)}$ - k-й вектор значений числа серий $\vartheta_z^{(k)}$ длиной Z элементов, $k=\overline{1,K}$:

$\Theta^{(k)}=\{\vartheta_1^{(k)},\ldots,\vartheta_z^{(k)},\ldots,\vartheta_Z^{(k)}\}$;

K - количество векторов $\Theta^{(k)}$;

$N\left(\vartheta_z^{(k)},\,\Theta^{(k)}\right)$ - код-номер двоичной последовательности, полученный для z-ой допустимой зоны по числу серий $\vartheta_z^{(k)}$ для вектора $\Theta^{(k)}$;

$N\left(\Theta^{(k)}\right)$ - значение кода-номера двоичной последовательности, полученное с учетом обработки всех Z допустимых зон для k-го вектора значений величин $\vartheta_z^{(k)}$.

Доказательство. Система полиадических ограничений делит всю двоичную последовательность на подпоследовательности разделенные нулевыми значениями, для которых $\lambda_i=1$. Обозначим число допустимых зон единиц через Z. Тогда для каждой z-ой зоны будет свое максимально возможное число серий единиц $0\le\vartheta_z^{(k)}\le\vartheta(max)_z^{(k)}$. При чем по условию теоремы будет выполняться равенство

$$\vartheta=\sum_{z=1}^{Z}\vartheta_z^{(k)}. \quad (8)$$

Значит количество векторов K может быть большим единицы и равно количеству различных комбинаций величин $\vartheta_z^{(k)}$, сумма которых равна ϑ. Каждая комбинация величин $\vartheta_z^{(k)}$ формирует множество $\Psi\left(\Theta^{(k)}\right)$ (долевое БП множество), являющееся частью множества $\Psi(m,\Lambda,\vartheta)$. При этом для объединения всех множеств $\Psi\left(\Theta^{(k)}\right)$ от 1 до K выполняется равенство

$$\Psi(m,\Lambda,\vartheta)=\bigcup_{k=1}^{K}\Psi\left(\Theta^{(k)}\right). \quad (9)$$

Рассмотрим формирование кода-номера $N(\Theta^{(k)})$ в пределах долевого БП множества. Поскольку на выбор конкретной последовательности $\vartheta_1^{(k)},\ldots,\vartheta_z^{(k)},\ldots,\vartheta_Z^{(k)}$ не существует никаких ограничений кроме (8), то формирование кода $N(\Theta^{(k)})$ проводится по каждой z-ой зоне независимо друг от друга. Следовательно, выполняется равенство

$$\Psi\left(\Theta^{(k)}\right)=\prod_{z=1}^{Z}\Psi\left(\Theta^{(k)},\vartheta_z^{(k)}\right), \quad (10)$$

где $\Psi\left(\Theta^{(k)},\vartheta_z^{(k)}\right)$ - подмножество множества $\Psi\left(\Theta^{(k)}\right)$ для z-ой зоны двоичной последовательности.

Для случая $z=2$ и значений кодов $N\left(\vartheta_1^{(k)},\,\Theta^{(k)}\right)=N_1$ и $N\left(\vartheta_2^{(k)},\,\Theta^{(k)}\right)=N_2$ обрабатываемой последовательности в соответствии с выражением (10) и принципа нумерации [3 - 6] в долевом БП множестве будет предшествовать $N_1\times N_2$ двоичных последовательностей. Обобщая данный результат для произвольного z получим выражение для вычисления значения кода-номера $N(\Theta^{(k)})$:

$$N(\Theta^{(k)})=\prod_{z=1}^{Z} N\left(\vartheta_z^{(k)},\,\Theta^{(k)}\right). \quad (11)$$

Рассмотрим вычисление кода-номера $N(m,\Lambda,\vartheta)$ для известных значений кодов $N(\Theta^{(k)})$. В соответствии с формулой (9) множества $\Psi(\Theta^{(k)})$ являются взаимонезависимыми. Тогда согласно комбинаторного правила сложения [5, 6] значение кода $N(m,\Lambda,\vartheta)$ для биноминально – полиадического числа находится как сумма значений кодов-номеров долевых БП множеств. ***Теорема доказана***.

Таким образом, полученная система выражений (6), (7) позволяет сформировать код-номер для двоичной последовательности и в силу неравенств (4), (5) его значение не будет превышать значений, полученных для полиадической и биноминальной нумерации. Это приводит к уменьшению затрат разрядов на представление кода-номера для биноминально - полиадического числа, а, следовательно, и к повышению степени сжатия.

Для дальнейшего увеличения степени сжатия данных необходимо рассмотреть варианты наложения дополнительных структурных характеристик. Исходя из анализа выражений (6) и (7), в качестве такой характеристики может выступать число (сумма) единиц в двоичных сериях.

Разработка метода двухпризнаковой биноминально – полиадической нумерации

Сформулируем определения двухпризнаковой БП нумерации.

Определение 2. Двухпризнаковой биноминально – полиадической нумерацией двоичных данных называется такая нумерация, которая формирует код-номер для двухпризнакового биноминального числа [4] в двоичном полиадическом пространстве.

Для нумерации двухпризнаковых БП чисел сформулируем и докажем теорему.

Теорема о двухпризнаковой биноминально – полиадической нумерации. Для двоичного биноминально полиадического числа $A=\{a_i\}_{i=\overline{1,m}}$ с дополнительным ограничением на число единиц χ можно поставить в соответствие код-номер $N(m,\Lambda,\vartheta,\chi)$, которые находится по формулам:

$$N(m,\Lambda,\vartheta)=\sum_{k=1}^{K_\chi}\sum_{u=1}^{U}\prod_{z=1}^{Z}N\left(\chi_{z,u}^{(k)},\,\vartheta_z^{(k)},\,\Theta^{(k)}\right); \qquad (12)$$

$$N\left(\chi_{z,u}^{(k)},\,\vartheta_z^{(k)},\,\Theta^{(k)}\right)=\sum_{i=1}^{m_z}a_i\,V(\chi_{z,u}^{(k)},\,\vartheta_z^{(k)})_i\ ; \qquad (13)$$

$$V(\chi_{z,u}^{(k)},\,\vartheta_z^{(k)})_i=\begin{pmatrix}\chi_{z,u,i-1}^{(k)}-1\\ \vartheta_{z,i-1}^{(k)}-1\end{pmatrix}\begin{pmatrix}m-i-\chi_{z,u,i-1}^{(k)}+1\\ \vartheta_{z,i-1}^{(k)}\end{pmatrix}, \qquad (14)$$

где $\chi_{z,u,i-1}^{(k)}$, $\vartheta_{z,i-1}^{(k)}$ - соответственно значения количества единиц и числа серий на $(i-1)$-ом шаге кодирования двоичной последовательности; $V(\chi_{z,u}^{(k)},\,\vartheta_z^{(k)})_i$ - весовой коэффициент i-го элемента z-го разряда двухпризнакового БП числа; $N\left(\chi_{z,u}^{(k)},\,\vartheta_z^{(k)},\,\Theta^{(k)}\right)$ - код-номер z-го разряда двухпризнакового БП числа; K_χ - количество долевых двухпризнаковых БП множеств.

Доказательство. Доказательство данной теоремы проводится по аналогии с доказательством теоремы о биноминально – полиадической нумерации. Отличительный момент состоит в том, что согласно определению 2 долевое БП множество $\Psi(\Theta^{(k)})$ прореживается долевыми двухпризнаковыми БП множествами $\Psi\left(\Theta_u^{(k)}\right)$ по числу единиц в сериях. В этом случае для каждого допустимого набора числа серий $\Theta^{(k)}=\{\vartheta_1^{(k)},\ldots,\vartheta_z^{(k)},\ldots,\vartheta_Z^{(k)}\}$ образуются наборы $\Theta_u^{(k)}$ чисел единиц

$$\Theta_u^{(k)}=\{\chi_{1,u}^{(k)},\ldots,\chi_{z,u}^{(k)},\ldots,\chi_{Z,u}^{(k)}\},$$

где $\mathbf{u}$ изменяется от $\mathbf{1}$ до $\mathbf{U}$, равному количеству комбинаций сформированных из величин $\chi_{z,u}^{(k)}$ с ограничениями:

$$\chi=\sum_{z=1}^{Z}\chi_{z,u}^{(k)}\,; \tag{15}$$

$$\vartheta_z^{(k)} \le \chi_{z,u}^{(k)} \le m_z-\vartheta_z^{(k)}+1\,. \tag{16}$$

Тогда, поскольку на выбор последовательности величин $\chi_{z,u}^{(k)}$ не накладывается других ограничений, то мощность множества $\Psi\left(\Theta_u^{(k)}\right)$ равна

$$\Psi\left(\Theta_u^{(k)}\right)=\bigcup_{u=1}^{U}\prod_{z=1}^{Z}\Psi\left(\Theta^{(k)},\vartheta_z^{(k)},\chi_{z,u}^{(k)}\right), \tag{17}$$

где $\Psi\left(\Theta^{(k)},\vartheta_z^{(k)},\chi_{z,u}^{(k)}\right)$ - подмножество долевого двухпризнакового БП множества для $\mathbf{z}$-ой зоны и $\mathbf{u}$-ой последовательности $\Theta_u^{(k)}$.

Отсюда с учетом формулы (17) код-номер $N\left(\Theta_u^{(k)}\right)$ долевого двухпризнакового БП числа в пределах множества $\Psi\left(\Theta_u^{(k)}\right)$ равен

$$N\left(\Theta_u^{(k)}\right)=\sum_{u=1}^{U}\prod_{z=1}^{Z}N\left(\Theta^{(k)},\vartheta_z^{(k)},\chi_{z,u}^{(k)}\right). \tag{18}$$

Для определения значения кода $N(m,\Lambda,\vartheta,\chi)$ по известным долевым кодам-номерам $N\left(\Theta_u^{(k)}\right)$ рассмотрим формирование множества $\Psi(m,\Lambda,\vartheta,\chi)$. Поскольку по определению двухпризнакового БП числа множества $\Psi\left(\Theta_u^{(k)}\right)$ формируются независимо друг от друга и полностью описывают допустимые классы двоичных последовательностей, то

$$\Psi(m,\Lambda,\vartheta,\chi)=\bigcup_{k=1}^{K_\chi}\Psi\left(\Theta_u^{(k)}\right). \tag{19}$$

Значит в соответствии с комбинаторным правилом сложения значение [5, 6] кода-номера $N(m,\Lambda,\vartheta,\chi)$ равно сумме отдельных долевых кодов $N\left(\Theta_u^{(k)}\right)$. ***Теорема доказана***.

Из доказанной теоремы вытекает следствие.

Следствие. Из анализа ограничений (15), (16) следует, что:

- мощность долевого двухпризнакового БП множества будет меньше мощности множества $\Psi\left(\Theta^{(k)}\right)$. Это обусловлено уменьшением допустимого множества значений количества единиц для заданного числа серий;

- количество множеств $\Psi\left(\Theta_u^{(k)}\right)$ будет меньше, чем количество множеств $\Psi\left(\Theta^{(k)}\right)$, т.е. $K_\chi \le K$. Это вызвано тем, что не все комбинации чисел серий удовлетворяют условиям (15), (16).

Поэтому за счет перехода от БП нумерации к двухпризнаковой БП нумерации достигается дополнительное уменьшение количества разрядов, отводимое на представление кода-номера двоичной последовательности.

Замечание. Для снижения количества операций для двухпризнаковой БП нумерации предлагается значения величин $\chi_{z,u,i-1}^{(k)}$, $\vartheta_{z,i-1}^{(k)}$ находить рекуррентно по правилу [6]:

- если $a_i=0$, то $\chi_{z,u,i}^{(k)}=\chi_{z,u,i-1}^{(k)}$ и $\vartheta_{z,i}^{(k)}=\vartheta_{z,i-1}^{(k)}$;
- если $a_i=a_{i+1}=1$, то $\chi_{z,u,i}^{(k)}=\chi_{z,u,i-1}^{(k)}-1$ и $\vartheta_{z,i}^{(k)}=\vartheta_{z,i-1}^{(k)}$;
- если $a_i=1$, а $a_{i+1}=0$, то $\chi_{z,u,i}^{(k)}=\chi_{z,u,i-1}^{(k)}-1$ и $\vartheta_{z,i}^{(k)}=\vartheta_{z,i-1}^{(k)}-1$.

Начальные значения признаков соответственно равны $\chi_0=\chi$ и $\vartheta_0=\vartheta$.

Кроме того, из анализа выражений (12), (13) следует, что вычисление весовых коэффициентов

$V(\chi_{z,u}^{(k)}, \vartheta_z^{(k)})_i$ не зависит от индексов зоны z, выборки u и индекса долевого двухпризнакового БП множества. Следовательно, дальнейшего снижение временных затрат для осуществления двухпризнаковой БП нумерации можно добиться путем организации конвейерной и параллельной обработки.

Оценка эффективности

Оценка эффективности методов биноминально – полиадической нумерации проводилась на основе обработки элементов полноцветных реалистических изображений. Эксперименты показали, что для изображений с различной степенью насыщенности дополнительный коэффициент сжатия за счет однопризнаковой и двухпризнаковой БП нумерации составил соответственно **1,9** и **3,1** раз.

Заключение

Таким образом, можно сделать выводы:

1. Разработана биноминально – полиадическая нумерация двоичных данных на основе формирования кода-номера двоичным последовательностям, удовлетворяющим одновременно биноминальным и полиадическим ограничениям. Сокращение исходного биноминального пространства достигается в результате отбора только тех биноминальных чисел, которые не имеют единичных элементов на запрещенных позициях. Наименьшее значение дополнительного коэффициента сжатия для $\mathbf{m = 8}$ составляет **2** раза.

2. Разработан метод двухпризнаковой биноминально – полиадической нумерации. В отличии от однопризнакового БП множества на двухпризнаковое БП множество накладываются дополнительные ограничения на число единиц в двоичной серии. Это приводит к дополнительному уменьшению мощности БП множества. При этом выигрыш по сжатию для двухпризнаковой БП нумерации относительно однопризнаковой БП нумерации достигает не менее **1,7** раза.

Литература

1. Зубарев Ю.В., Дворкович В.П. Цифровая обработка телевизионных и компьютерных изображений М.: Международный центр научной и технической информации, 1997. – 212 с.

2. Ватолин В.И., Ратушняк А., Смирнов М., Юкин В. Методы сжатия данных. Устройство архиваторов, сжатие изображений и видео. – М.: ДИАЛОГ – МИФИ, 2002. – 384 с.

3. Королев А.В. Версификационная избыточность изображений // Інформаційно - керуючи системи на залізничному транспорті. – 2002. – № 2. – С. 26 – 30.

4. Королев А.В. Полиадическое кодирование одноцветных областей изображений по числу двоичных серий // Інформаційно - керуючи системи на залізничному транспорті. – 2002. – № 1. – С. 3 - 9.

5. Королев А.В. Баранник В.В. Оценка количества информации изображения по числу серий одинаковых элементов // Системи обробки інформації. – Харків: НАНУ, ПАНМ, ХВУ. – 2002. – Вип. 2(18).– С. 43 – 46.

6. Баранник В.В. Метод двухпризнакового биноминального кодирования двоичных данных // Системи обробки інформації. – Харків: НАНУ, ПАНМ, ХВУ. – 2002. – Вип. 6(22).– С. 24 – 28.

Поступила в редакцию 11.09.03

Рецензент: доктор техн. наук, проф. Бильчук В.М., Харьковский военный университет