

**Аспірант Проскурін Д.В.**

**Національний технічний університету України  
«Київський політехнічний інститут»**

## **ЗАВАДОСТІЙКЕ КОДУВАННЯ ДАНИХ ПРИ ЇХ ПОДАННІ У ВИГЛЯДІ ДВОВИМІРНОГО ГРАФІЧНОГО КОДУ**

### **Вступ**

У системах автоматичної ідентифікації широко використовується графічне кодування даних [1]. Дедалі більшого поширення набувають двовимірні графічні коди (ГК), в яких інформація подається у вигляді чорно-білих елементів (чорний елемент позначає двійкову одиницю, білий – нуль), що розміщуються по горизонталі та вертикалі (рис. 1). Інформаційна ємність двовимірних ГК, як правило, становить декілька сотень алфавітно-цифрових символів [2].

Головною перевагою використання ГК є їх здатність до автоматичного (машинного) зчитування.

### **Постановка задачі**

Графічний код розміщують на поверхні об'єкта обліку. Оскільки ГК є відкритим носієм інформації, що може зазнавати зовнішніх впливів, в тому числі й механічних ушкоджень, то важливою задачею є забезпечення надійного зберігання графічно-кодової інформації. Досягти цього можна лише за рахунок використання потужних завадостійких кодів, здатних виправляти багатократні спотворення.

Метою статті є розроблення методики подання даних у графічно-кодованому вигляді з підвищеними показниками надійності.

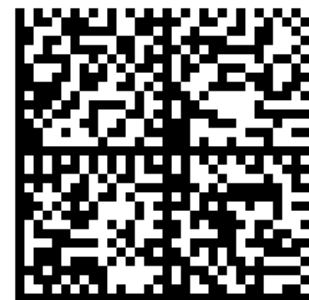


Рис. 1. Загальний вигляд двовимірного графічного коду

### **Кодування вхідних даних**

Процес кодування вхідних символних даних полягає в перетворенні їх у числову форму. Для кодування довільних вхідних символних

послідовностей використовують декілька режимів перетворення: текстовий режим, у якому кожним трьом суміжним символам  $C_1C_2C_3$  вхідного потоку ставлять у відповідність два числа  $CW_1, CW_2$ , які називають кодословами і обчислюють за формулами

$$\begin{aligned} CW_1 &= (1600 \cdot C_1 + 40 \cdot C_2 + C_3 + 1) \operatorname{div} 256, \\ CW_2 &= (1600 \cdot C_1 + 40 \cdot C_2 + C_3 + 1) \operatorname{mod} 256, \\ CW_1, CW_2 &\in \{0, 1, 2, \dots, 255\}; \end{aligned}$$

режим ASCII, в якому чотирьом суміжним символам ставлять у відповідність 3 кодослова; цифровий режим, у якому сімом суміжним десятковим цифрам ставлять у відповідність 4 кодослова.

Результатом кодування вхідних символьних даних є послідовність кодослів.

Наприклад, повідомленню

01234 Київ а/с 12 РЕКОМЕНДОВАНИЙ ЛИСТ, вага 0.125, вартість 6.79

відповідає послідовність з 52-х кодослів

$$\begin{array}{l} 131\ 153\ 52\ 79\ 239\ 123\ 44\ 56\ 188\ 127\ 254\ 142\ 230\ 12\ 146\ 201 \\ 34\ 211\ 32\ 10\ 120\ 144\ 151\ 23\ 110\ 98\ 219\ 34\ 77\ 121\ 71\ 46\ 152\ 239 \\ 34\ 121\ 254\ 32\ 48\ 142\ 53\ 239\ 45\ 117\ 12\ 210\ 134\ 55\ 254\ 54\ 46\ 209 \end{array} \quad (1)$$

### Кодування даних за допомогою коректувального коду

Для забезпечення високої надійності зберігання даних у формі ГК необхідно використовувати коректувальний код, здатний виправляти багатократні спотворення двох видів – помилки та стирання [3].

Стирання – це таке ушкоджене кодослово, значення якого не належить полю Галуа  $GF(256)$  (тобто, множині значень  $0, 1, 2, \dots, 255$ ). Фізично стирання – це символ, який неможливо відсканувати. Тобто, на виході сканера буде отримано значення, яке неможливо ототожнити з жодним зі значень  $0, 1, 2, \dots, 255$ . Місцезнаходження таких значень у зчитаному слові вказує на місця виникнення стирань. Причиною появи стирань можуть бути механічні ушкодження ГК-позначок або дефекти друку.

Помилка являє собою перехід правильного кодослова в інше (неправильне) кодослово, але таке, що належить полю Галуа  $GF(256)$ . Значенням помилки може бути будь-яка величина:  $1, 2, \dots, 255$ . На відміну від стирань, місцезнаходження помилок перед етапом декодування невідомі.

Властивість виправляти помилки і стирання має коректувальний код Ріда-Соломона [3, 4].

Якщо довжина блоку коду Ріда-Соломона (загальна кількість кодослів) дорівнює  $n$ , то  $n = k + r$ , де  $k$  – кількість інформаційних кодослів, а  $r$  – кількість контрольних кодослів.

Для обчислення контрольних кодослів використовується твірний поліном  $g(x)$  степеня  $r$ , де  $r$  – кількість контрольних кодослів.

Твірний поліном  $g(x)$  обчислюють наступним чином:

$$g(x) = (x+2^1)(x+2^2)\dots(x+2^r).$$

Арифметичні операції над коефіцієнтами полінома виконують в полі GF(256). Елемент поля можна подати у вигляді многочлена. Наприклад, елементу 68 відповідає многочлен  $x^6 + x^2$ , коефіцієнтами якого є розряди 8-розрядного двійкового подання числа 68 ( $68 = 01000100$ ).

Операцію додавання елементів поля GF( $2^8$ ) можна виконувати як додавання многочленів (коефіцієнти при однакових степенях  $x$  додають за модулем 2).

Операцію множення елементів поля GF( $2^8$ ) виконують за модулем незвідного многочлена степеня 8:  $m(x) = x^8 + x^5 + x^3 + x^2 + 1$ .

Щоб отримати добуток елементів поля, необхідно виконати множення відповідних многочленів (коефіцієнти при однакових степенях  $x$  слід додавати за модулем 2), а отриманий многочлен поділити на незвідний многочлен  $m(x)$ . Результатом буде остача від ділення (частку слід відкинути).

При діленні отриманого многочлена на незвідний многочлен операція віднімання замінюється на операцію додавання за модулем 2.

Операцію кодування можна описати рівнянням

$$c(x) = x^r \cdot a(x) + z(x),$$

де  $a(x)$  – інформаційний многочлен степеня  $k-1$

$$a(x) = \sum_{i=0}^{k-1} a_i x^i = a_{k-1} x^{k-1} + \dots + a_1 x + a_0,$$

якому відповідає послідовність інформаційних кодовекторів

$$\mathbf{A} = (a_{k-1}, a_{k-2}, \dots, a_2, a_1, a_0), \quad a_i \in \{0, 1, 2, \dots, 255\},$$

а  $z(x)$  є многочленом остачі від ділення  $x^r \cdot a(x)$  на твірний многочлен  $g(x)$ .

Многочлену остачі відповідає вектор остачі

$$\mathbf{Z} = (z_{r-1}, z_{r-2}, \dots, z_1, z_0).$$

Вектор  $\mathbf{C}$ , якому відповідає многочлен  $c(x)$ , подамо у вигляді:

$$\mathbf{C} = (a_{k-1}, a_{k-2}, \dots, a_1, a_0, z_{r-1}, z_{r-2}, \dots, z_1, z_0).$$

Закодуємо послідовність кодослів (1) кодом Ріда-Соломона з параметром  $r = 36$ .

Результатом кодування є послідовність з 98-и кодослів, у якій перші 62 кодослова є інформаційними (1), а решта 32 – контрольними:

$$\begin{array}{l} 32 \ 110 \ 57 \ 211 \ 43 \ 212 \ 75 \ 103 \ 46 \ 122 \ 41 \ 58 \ 202 \ 194 \ 133 \ 151 \\ 115 \ 43 \ 29 \ 242 \ 44 \ 50 \ 223 \ 32 \ 107 \ 89 \ 30 \ 196 \ 137 \ 58 \ 129 \ 48 \end{array} \quad (2)$$

Послідовність, що складається з 98-и кодослів (об'єднання послідовностей (1) і (2)), подається у двійковій формі (кодослову відповідає байт – 8 двійкових розрядів). Ставлячи у відповідність

одинацям чорні елементи, а нулям – білі, отримаємо ГК, зображений на рис. 1.

Таким чином, методика подання даних у вигляді ГК з підвищеним показником надійності полягає в наступному.

1. Перетворити символи вхідної послідовності у числові значення (кодослова).

2. На основі отриманих інформаційних кодослів обчислити контрольні кодослова та долучити їх до інформаційних, в результаті чого утвориться блок коду Ріда-Соломона.

3. Кодослова блоку Ріда-Соломона подати у двійковому вигляді та утворити бітову матрицю. Замінивши в ній нулі та одиниці відповідними білими та чорними елементами, отримати ГК-позначку двовимірного графічного коду.

## Висновки

Застосування коректувального коду Ріда-Соломона є ефективним засобом підвищення надійності двовимірних графічних кодів.

Довжина блоку коду Ріда-Соломона залежить від степеня незвідного многочлена  $t(x)$ : якщо степінь  $t(x)$  дорівнює 8, то максимальна довжина блока становить 255 кодослів; якщо обрати незвідний многочлен степеня 10, то довжина блоку Ріда-Соломона зросте до 1023. Відповідно, зросте ємність графічно-кової позначки.

Подальші дослідження слід зосередити на з'ясуванні залежності ущільнення вхідних даних від величини степеня незвідного многочлена.

## Література

1. Дичка І.А. Зберігання інформації у вигляді багатокольорових штрихових кодів та їх обробка. – К.: ІВЦ “Видавництво “Політехніка”, 2003.
2. Тарасенко В.П., Дичка І.А., Голуб В.І. Штрихові коди та їх застосування. – К.: “Корнійчук”, 2000.
3. Р. Блейхут. Теория и практика кодов, контролирующих ошибки. – М.: “Мир”, 1986.
4. *International Standard. Information technology – International symbology specification – Data Matrix. ISO/IEC 16022:2000(E).*