



Н. Я. Возна, А. І. Сидор

Тернопільський національний економічний університет, м. Тернопіль, Україна

КРИТЕРІЙ СТРУКТУРНОЇ СКЛАДНОСТІ ТА ОЦІНКИ РОЗШИРЕНОЇ ХЕММІНГОВОЇ ВІДДАЛІ ДЛЯ ПРОСТОРОВИХ ОБ'ЄКТІВ

Охарактеризовано сучасний стан вирішення проблеми розпізнавання образів на основі теорії Хеммінгової віддалі. Викладено теоретичні основи методів розпізнавання образів. Наведено недоліки відомих методів розпізнавання образів у Хеммінговому просторі. Проаналізовано методи оцінки структурної складності поліфункціональних даних, запропонований критерій та система атрибутів формалізації структуризованих даних. Наведено розрахунки розширеної оцінки Хеммінгової віддалі для просторових об'єктів. Наведено формалізацію відомих методів розпізнавання одновимірних образів. Наведено характеристики ансамблів для сукупностей сигналів оптимальних сигнальних просторів. Наведено вирішення задачі оцінки структурної складності поліфункціональних даних на основі вагових коефіцієнтів оцінок інформативності. Представлено кількісну, оптимальну оцінку функціонально-структурної складності компонентів складної системи у вигляді відношення. Реалізовано методи розпізнавання образів на основі оцінки структурної складності. Запропоновано способи кодування атрибутів та розрахунку Хеммінгової віддалі на прикладі символіки гральних карт. Наведено порівняльні діаграми оцінок Хеммінгової віддалі для пар символів, що використовуються для різних методів знаходження Хеммінгової віддалі. Охарактеризовано недоліки і переваги різних запропонованих способів кодування атрибутів зображення.

Ключові слова: сигнальний простір; Евклідова відстань; розпізнавання образів; інформаційні технології; кодування даних.

Вступ. Теорію розпізнавання образів широко використовують у різних галузях знань, інформаційних та комп'ютерних системах, що застосовуються в біології, медицині, астрономії, навігації і т.ін.

Такі задачі успішно розв'язують за допомогою розробки програмно-апаратних засобів оцінки розширеної Хеммінгової віддалі. Незважаючи на широке застосування теорії Хеммінгового простору у різних галузях науки та техніки, актуальною задачею є розроблення критеріїв структурної складності певних класів образів та розроблення методів визначення Хеммінгової віддалі на основі різницевої віддалі модульних одиниць структурної складності образів. Такий підхід дає змогу зменшити об'єм обчислень визначення Хеммінгової віддалі між окремими компонентами образів з різною оцінкою структурної складності образів.

Теорія розпізнавання образів базується на оцінках Евклідової відстані між i -м і j -м об'єктами (тобто мірою схожості, близькості об'єктів між собою за всією сукупністю використовуваних ознак), метрики Мінковського, що фактично є великим сімейством метрик і містить у собі звичайну Евклідову відстань, манхеттенську відстань, яка є сумою модулів різниць відповідних ознак об'єктів, метрики "домінування" або Sup-метрики. Досить важливою є також так звана відстань Махаланобіса, яка має специфічні властивості (Krulikovsky, et al., 2016; Nykolaichuk & Zavediuk, 2010). У разі використання дихотомічних (що мають всього два значення) якісних ознак образів широко використовується зважена відстань Хеммінга.

Відомі методи вузькоспеціалізовані і не характеризуються універсальністю до різних образів. Кореляційні методи найефективніші, але вважають низькошвидкісними та апаратно-складними, оскільки передбачають складне попереднє статистичне оброблення образів, зокрема обчислення математичного сподівання, дисперсії, автокореляційної функції. Відомі методи розпізнавання в Хеммінговому просторі недостатньо ефективні, оскільки використовуються для образів, які описуються бінарними векторами, а також не враховують можливість застосування різних аналітичних виразів взаємодіяючих функцій та можливостей кодування даних у різних теоретико-числових базисах (ТЧБ).

Досвід досліджень наукової школи Я. М. Николаїчука, які охоплюють успішне вирішення задач істотного спрощення алгоритмів, зменшення апаратної складності та підвищення швидкодії цифрових процесорів кореляційного опрацювання сигналів, створює позитивні умови застосування методів кореляційного розпізнавання образів для багатьох класів джерел інформації (Nykolaichuk, 2010; Nykolaichuk, 2012; Vozna, et al., 2013).

Мета дослідження полягає в розробленні критеріїв структурної складності певних класів образів та методів визначення Хеммінгової віддалі на основі різницевої віддалі модульних одиниць структурної складності образів.

Методи розпізнавання образів. Класично, виділяють такі методи розпізнавання образів (Krulikovsky, et al., 2016):

Цитування за ДСТУ: Возна Н. Я., Сидор А. І. Критерій структурної складності та оцінки розширеної Хеммінгової віддалі для просторових об'єктів. Науковий вісник НЛТУ України. 2017. Вип. 27(4). С. 159–165.

Citation APA: Vozna, N. Ya, & Sydor, A. I. (2017). Criterion of Structural Complexity and Evaluation of Extended Hemming Distance for Spatial Objects. Scientific Bulletin of UNFU, 27(4), 159–165. <https://doi.org/10.15421/40270434>

- перебору, коли проводиться порівняння з базою даних, де для кожного виду об'єктів представлені всі можливі модифікації відображення;
- аналізу та диференціації компонентів образу;
- використання штучних нейронних мереж;
- використання генетичних алгоритмів.

Метод перебору характеризується найбільшою обчислювальною та часовою складністю і може застосовуватись в тому випадку, коли число класів образу є скінчене, тобто не перевищує заданого числа.

Метод аналізу та диференціації компонентів образу характеризується поліноміальною складністю, базується на теорії моделювання в байєсівському підході та в застосуванні кодових віддалей Хеммінгового простору.

Метод використання штучних нейронних мереж базується на принципах навчання та рекурсивної асимптотичної складності. Методи на основі використання генетичних алгоритмів базуються на принципах еволюційного пошуку й поєднують комп'ютерне моделювання генетичних процесів у природних і штучних системах (Krulikovsky, et al., 2016).

Важливим компонентом методів та інформаційних технологій розпізнавання образів є попереднє їх опрацювання та перетворення вхідних даних, до якого належать такі методи: алгебраїчний; кореляційний; структурний; спектральний; геометричний; на базі нечітких множин; статистичний; просторово-зв'язаних процесів.

Серед зазначених методів перспективу має структурний метод, який успішно застосовується в теорії мереж Петрі та під час оцінки структури матричних моделей руху даних (Nykolaichuk, Vozna, & Pitukh, 2013). Окрім цих застосувань, доцільно розширити сферу застосування цього методу в галузі розпізнавання образів. Але для цього потрібно розвивати критерії та теорію оцінок структурної складності. Формалізація відомих методів розпізнавання одновимірних образів описується такими функціоналами (табл. 1) (Krulikovsky, et al., 2016).

Табл. 1. Формалізація методів розпізнавання одновимірних образів

№	Формула
1	$d_{ij} = \left(\sum_{k=1}^z (x_{ik} - x_{jk})^2 \right)^{1/2}$
2	$d_{ij} = \left(\sum_{k=1}^z x_{ik} - x_{jk} ^p \right)^{1/p}$
3	$d_{ij} = \left(\sum_{k=1}^z x_{ik} - x_{jk} ^\infty \right)^{1/\infty}$
4	$d_{ij} = (X_i - X_j)^T S^{-1} (X_i - X_j)$
5	$d_{ij} = \sum_{k=1}^z x_{ik} - x_{jk} $

У цій таблиці введено такі позначення: d_{ij} – значення Евклідової відстані між i -м і j -м об'єктами; x_{ik} – чисельне значення k -тої ознаки для i -го об'єкта; x_{jk} – чисельне значення k -тої ознаки для j -го об'єкта; z – кількість ознак, якими описуються об'єкти.

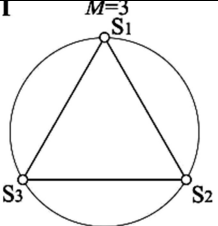
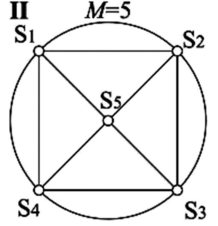
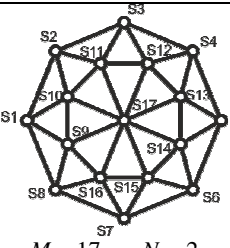
Розрахунки розширеної оцінки Хеммінгової віддалі. Розширена Хеммінгова віддаль може застосовуватись в різноманітних галузях знань, до яких відносять такі теорії: сигналів, інформації, випадкових процесів, кодування моделей джерел інформації та руху даних, графічних образів, а також аудіо- і відеозображень та схемотехніки.

Аналогічні задачі розрахунку розширеної оцінки Хеммінгової віддалі існують в джерелах інформації, які реалізуються двовимірними масивами даних 2D, тривимірними масивами даних 3D, які представляють розв'язки томографії, голограм, кристалічних решіток та ін.

Відома загальна формула верхньої кодової віддалі між двомірними сигналами, яка має вигляд: $d_{ij} = \alpha \sqrt{\log_2 M} / \sum E_i$, де: M – кількість сигналів, E – енергія кожного сигналу, d – лінійна кодова віддаль у сигнальному просторі.

Приклади оптимальних сигнальних просторів для сукупностей сигналів $M = 2^k$ та $M = 2^k + 1$ наведено у табл. 2.

Табл. 2. Характеристики ансамблів з числом $M=3,5,9,17$

Двовірні простори з числом сигналів $M = 3,5,9$	d_{\min} d_{\max} d_c	$E_{1,2,\dots}$ E_m E_c	α_c ν_c	Двовірні простори з числом сигналів $M = 17$	d_{\min} d_{\max} d_c	$E_{1,2,\dots}$ E_m E_c	α_c ν_c
1	2	3	4	5	6	7	8
I  $M=3$	1,732	$E_{1,2,3}=1,0$ $E_c=1,0$ $E_m=1,0$	0,662 0,5	–	0,434 2,0 1,217	$E_{1,8}=1,0$ $E_{9-16}=0,565$ $E_{17}=0$ $E_c=0,737$ $E_m=1,0$	0,719 3,087
II  $M=5$	1,0 2,0 1,365	$E_{1-4}=1,0$ $E_5=0,0$ $E_c=0,8$ $E_m=1,0$	0,64 1,332	 $M = 17$ $N = 2$	0,482 2,0 1,241	$E_{1,8}=1,0$ $E_{9-16}=0,63$ $E_{17}=0$ $E_c=0,767$ $E_m=1,0$	0,716 3,087

1	2	3	4	5	6	7	8
III <i>M</i> =9	0,707	$E_{1,3}=0,707$ $E_{5,7}=0,707$ $E_{2,4,6,8}=1,0$ $E_9=0,0$ $E_c=0,753$ $E_m=1,0$	0,591 2,127	 <i>M</i> = 17 <i>N</i> = 2	0,5 2,0 1,25	$E_{1-10} = 1,0$ $E_{11-16} = 0,5$ $E_{17} = 0$ $E_c = 0,765$ $E_m = 1,0$	0,723 3,087
IV <i>M</i> =9	0,75 2,0 1,164	$E_{1-8}=1,0$ $E_9=0,0$ $E_c=0,889$ $E_m=1,0$	0,58 2,127	 <i>M</i> = 17 <i>N</i> = 2	0,518 2,0 1,259	$E_{1-12} = 1,0$ $E_{13-17} = 0,46$ $E_c = 0,841$ $E_m = 1,0$	0,694 3,087
V <i>M</i> =9	0,53 2,0 0,996	$E_{1-4}=1,0$ $E_{5-8}=0,53$ $E_9=0,0$ $E_c=0,63$ $E_m=1,0$	0,465 2,0	 <i>M</i> = 17 <i>N</i> = 2	0,465 2,0 1,233	$E_{1-10} = 1,0$ $E_{11-17} = 0,536$ $E_c = 0,809$ $E_m = 1,0$	0,693 3,087

Відомий метод оцінки структурної складності схем комп'ютерної мікроелектроніки на основі ентропійного підходу, який запропонував професор М. В. Черкаський (Cherkaskyi & Murad, 2004). Суть цього методу полягає у тому, що на основі досліджуваної структури будується матриця суміжності і обчислюється коефіцієнт структурної складності згідно з виразом:

$$S = -C \log_2 \frac{C}{r(r-1)},$$

де: *C* – кількість елементів матриці ідентичності системи, *r* – розмір матриці. Недоліком такого критерію оцінки структурної складності є невизначеність розра-

хунків під час аналізу регулярних та однорідних структур, коли виконується $\log_2 1$.

Розв'язок задачі оцінки структурної складності поліфункціональних даних (ПФД) на основі вагових коефіцієнтів оцінок інформативності викладено у роботах (Vozna, 2014; Vozna & Nykolaichuk, 2013), де проблему оцінки коефіцієнта структуризованості відображуваних ПФД вирішено таким чином. Експериментальним способом визначено коефіцієнт складності компонентів атрибутів поліфункціональних даних – α_i та присвоєно вагові коефіцієнти оцінок інформативності їх параметрів P_i (табл. 3):

Табл. 3. Вагові коефіцієнти оцінок інформативності

№ з/п	Познач. елемента	Зміст елемента	Символ	α_i
1	2	3	4	5
1	<i>l</i>	Лінія		1
				1,5
				1,1
				1,1
				1,2
				1,2
2	<i>P</i>	Поворот		2
				2,2
				2,2
3	<i>x</i>	Пересічення		3
				3,1
				3,1
4	<i>d</i>	Дотик		2
				2,2

1	2	3	4	5
5	<i>r</i>	Розгалуження		4 4,2 6,2
6	<i>h</i>	Спосіб заливки		2
7	<i>z</i>	Направлений зв'язок		2 3 2,4 2,5 3,4 3,5
8	<i>b</i>	Літера	Aa...Яя, ..., Aa...Яя, Aa...Яя, ..., Aa...Яя, Aa...Яя, ..., Aa...Яя, Aa...Яя, ..., Aa...Яя, Aa...Яя, ..., Aa...Яя Aa...Zz, ..., Aa...Zz Aa...Zz, ..., Aa...Zz Aa...Zz, ..., Aa...Zz Aa...Zz, ..., Aa...Zz	8-10
	<i>c</i>	Цифра	1, 2, ..., 0, ..., I, 2, ..., 0	4
	<i>i</i>	Індекс	1, 2, ..., 0, a, A, ..., 1, 2, ..., 0, a, A	4
	<i>s</i>	Символ	©, ®, π, ψ, ω, &, %, @, \$, Θ, №, Σ, J, ∞, ☉, ☼, ♪, ...	4
	<i>n</i>	Знак	+, -, <, >, =, ≠, ≈, ≠, ≤, ≥, (, ", {, !, ?, ;, :; ...	2

Враховуючи вказані показники, отримуємо коефіцієнт структурної складності $k_c = \sum_{i=1}^n \alpha_i P_i$.

Враховуючи також функціонально-інформативну характеристику ПФД f_i та їх суму $\sum_{j=1}^m f_j$, отримаємо кількісну, оптимальну оцінку функціонально-структурної складності представлення компонентів складної системи у вигляді відношення:

$$k_e = K \cdot \sum_{j=1}^m f_j / \sum_{i=1}^n \alpha_i P_i \Rightarrow \max, \quad (1)$$

K – ідентифікатор рівня ПФД ($K = n, \dots$ – відповідно для n -рівневих зображень). При цьому рівень структурізації ПФД підвищується за максимізації відношення (1). Реалізація методу розпізнавання образів на основі оцінки структурної складності.

Розглянемо реалізацію методу розпізнавання образів на основі оцінки структурної складності (СС) на прикладі символіки гральних карт. Під час розрахунку Хеммінгової оцінки СС для гральних карт базуємося на такій симетрії числа однотипних елементів, які кодуються двійковими логічними ознаками (рис. 1).

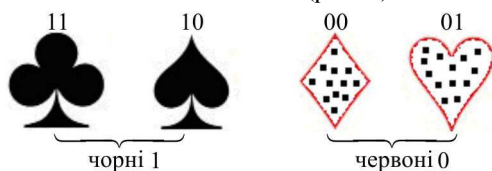


Рис. 1. Символічні базові зображення атрибутів гральних карт та їх двійкові коди

На рис. 2 запропоновано спосіб кодування символіки гральних карт, на основі яких можна показати Хеммінгову відстань між ними.

Окрім кольору (k) фігури і типу фігури, потрібно також закодувати кількість увігнутих (vv) та випуклих

(vp) елементів, які є компонентами досліджуваних символів, тобто:

$$\heartsuit \Rightarrow 2vv + 2vp; \spadesuit \Rightarrow 4vv; \clubsuit \Rightarrow 3vv + 3vp; \spadesuit \Rightarrow 5vv + 2vp.$$

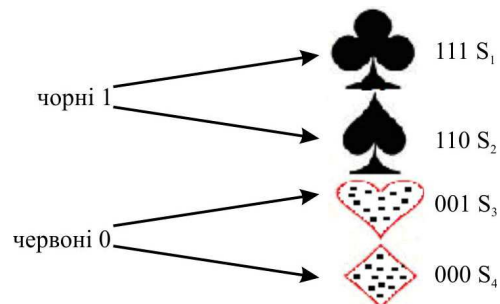


Рис. 2. Система кодування символіки гральних карт

Формула для визначення Хеммінгової віддалі для такого кодування матиме такий вигляд:

$$H_{dij} = H_k + H_{vv} + H_{vp}, \quad H_k = |k_1 - k_2|,$$

$$H_{vv} = |vv_1 - vv_2|, \quad H_{vp} = |vp_1 - vp_2|.$$

З урахуванням кодування, наведеного на рис. 1 та рис. 2 розрахуємо Хеммінгову відстань між незалежними парами символів (табл. 4).

Табл. 4. Перший метод визначення Хеммінгової віддалі

№	Пара символів	H_k	H_{vv}	H_{vp}	H_{dij}
1	♥-♦	0-0	2-4	2-0	4
2	♥-♣	0-1	2-3	2-3	3
3	♥-♠	0-1	2-5	2-2	4
4	♦-♣	0-1	4-3	0-3	5
5	♦-♠	0-1	4-5	0-2	4
6	♣-♠	1-1	3-5	3-2	3

Таке кодування досить часто буде давати однакову Хеммінгову відстань. Щоб уникнути цього, потрібно враховувати асиметрію і симетрію фігур. Якщо врахувати кількість елементів, за якими будуть відрізнятися вказані фігури, отримаємо оцінки Хеммінгової віддалі, які наведено у табл. 5.

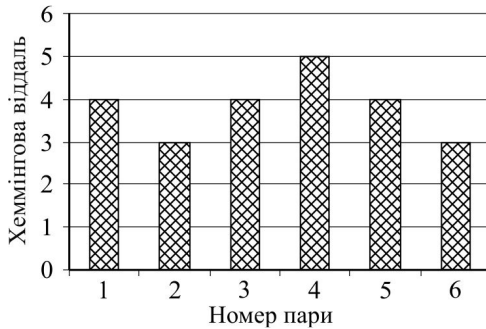


Рис. 3. Порівняльна діаграма оцінок Хеммінгової віддалі для пар символів, наведених у табл. 4

Табл. 5. Другий метод визначення Хеммінгової віддалі

№	Пара символів	H_k	H_{vv}	H_{vp}	H_{as}	H_{dij}
1	♥-♦	0-0	2-4	2-0	4	8
2	♥-♣	0-1	2-3	2-3	10	13
3	♥-♠	0-1	2-5	2-2	11	15
4	♦-♣	0-1	4-3	0-3	6	11
5	♦-♠	0-1	4-5	0-2	5	9
6	♣-♠	1-1	3-5	3-2	7	10

Таку оцінку розраховують за формулами:

$$H_{dij} = H_k + H_{vv} + H_{vp} + H_{as}, \quad H_k = |k_1 - k_2|,$$

$$H_{vv} = |vv_1 - vv_2|, \quad H_{vp} = |vp_1 - vp_2|,$$

$$H_{as} = vv_1 + vv_2 + vp_1 + vp_2 - sp,$$

де sp – спільні елементи пар символів.

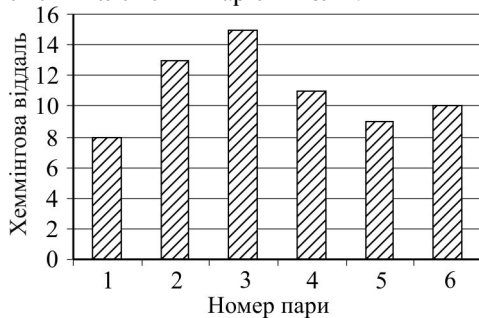


Рис. 4. Порівняльна діаграма оцінок Хеммінгової віддалі для пар символів, наведених у табл. 5

Точнішу Хеммінгову віддаль отримаємо між цими 4-ма класами зображень, маючи компоненти та два кольори □ червоний і ■ чорний, а також закодувавши увігнуті та опуклі елементи вказаних фігур. Закодуємо логічними символами: 1 – чорний колір (k); 0 – червоний; наявність елементів ввігнутих вниз ($vv-$) – символом 1 їх відсутність – символом 0; наявність увігнутих вгору ($vv+$) відповідно – 1, а їх відсутність відповідно – 0; відсутність випуклих елементів – 00, випуклі елементи (vp) вниз – 01, випуклі вгору – 10 і випуклі вбік – 11. Для вказаного на рис. 1 способу кодування елементів фігур отримаємо кодування, наведене в табл. 6.

Табл. 6. Кодування символіки гральних карт

Символ	vp	$vv+$	$vv-$	k
♥	01	1	0	0
♦	00	1	1	0
♣	11	0	1	1
♠	10	0	1	1

Згідно із вказаними атрибутами отримаємо оцінки Хеммінгової віддалі, які наведено в табл. 7.

З діаграми видно, що мінімальна Хеммінгова віддаль буде між ♣ і ♠, а максимальна – ♥ і ♠.

Доцільно використати більш просте кодування для вказаних фігур, за якого не буде враховуватися колір, а тільки кількість увігнутих і випуклих елементів.

Табл. 7. Третій метод визначення Хеммінгової віддалі

№	Пара символів	H_{vp}	H_{vv+}	H_{vv-}	H_k	$H_{dij} (2)$	$H_{dij} (10)$
1	♥-♦	01-00	1-1	0-1	0-0	01010	10
2	♥-♣	01-11	1-0	0-1	0-1	10111	23
3	♥-♠	01-10	1-0	0-1	0-1	11111	31
4	♦-♣	00-11	1-0	1-1	0-1	11101	29
5	♦-♠	00-10	1-0	1-1	0-1	10101	21
6	♣-♠	11-10	0-0	1-1	1-1	01000	8

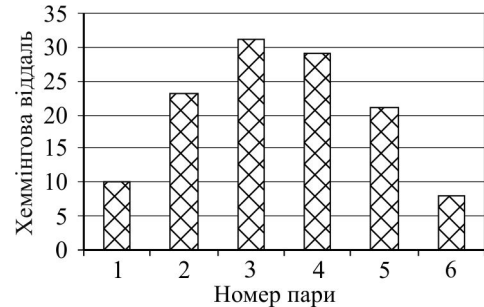


Рис. 5. Порівняльна діаграма оцінки Хеммінгової віддалі для пар символів, наведених у табл. 7

Позначимо через 00-2 vv , 01-3 vv , 10-4 vv , 11-5 vv , 00-0 vp , 01-2 vp , 11-3 vp , тоді коди елементів будуть такими: ♥ 0 0 0 1; ♦ 1 0 0 0; ♣ 0 1 1 1; ♠ 1 1 0 1. Згідно із вказаними атрибутами отримаємо Хеммінгову віддаль, наведену в табл. 8. З діаграми видно, що мінімальна Хеммінгова віддаль буде між ♦ і ♠, а максимальна – ♥ і ♠.

Табл. 8. Четвертий метод визначення Хеммінгової віддалі

№	Пара символів	H_{vv}	H_{vp}	$H_{dij} (2)$	$H_{dij} (10)$
1	♥-♦	00-10	01-00	1001	9
2	♥-♣	00-01	01-11	0110	6
3	♥-♠	00-11	01-01	1100	12
4	♦-♣	10-01	00-11	1111	15
5	♦-♠	10-11	00-01	0101	5
6	♣-♠	01-11	11-01	1010	10

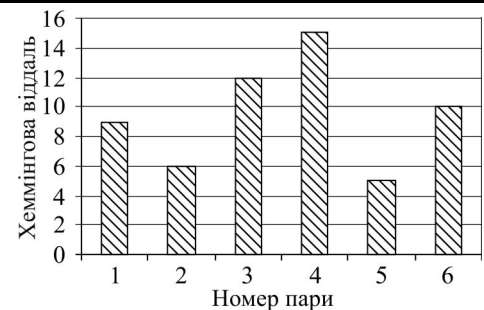


Рис. 6. Порівняльна діаграма оцінки Хеммінгової віддалі для пар символів, наведених у табл. 8

Використання 2 кольорів і 2 компонент за різних положень з урахуванням їх кількості, симетрії, асиметрії, кількість різних компонент структури складності по числу компонент в загальному надає набір ознак, за допомогою яких можна однозначно закодувати фігури, при порівнянні яких отримаємо різну Хеммінгову віддаль для однозначного розпізнавання.

На рис. 7 показано порівняльну діаграму запропонованих способів кодування атрибутів зображень символів гральних карт, з якої видно, що перший спосіб простий у кодуванні, однак він досить часто дає однакову Хеммінгову віддаль, що може призвести до помилкового розпізнавання образу.

Другий спосіб кодування дає добру Хеммінгову віддаль, однак він потребує додаткових обчислень для кожної пари порівнювальних символів. Третій спосіб також дає добрі оцінки Хеммінгової віддалі, проте для

його кодування потрібно використовувати більшу кількість пам'яті. Четвертий спосіб є простим у кодуванні і має добру Хеммінгову віддалю, проте потребує додаткового підрахунку кількості елементів.



Рис. 7. Порівняльна діаграма запропонованих методів визначення Хеммінгової віддалі

Висновок. Запропоновані способи кодування атрибутів та розрахунку Хеммінгової віддалі на прикладі символіки гральних карт є основою для застосування таких технологій визначення Хеммінгової віддалі для комп'ютерного розпізнавання дорожніх знаків, символіки хімічних сполук, показників маніпульованих сигналів систем, а також двовимірних зображень та структурних схем компонентів мікроелектроніки.

Перелік використаних джерел

Cherkaskyi, M. V., & Murad, Kh. Kh. (2004). Skladnist prystroiu keruvannia. *Visnyk Natsionalnoho universytetu "Lvivska politekhnika"*.

ka". Series: Komp'uterna inzheneriia ta informatsiini tekhnolohii, 521, 3–7. [in Ukrainian].
 Krulikovskyy, B. B., Sydor, A. I., Zastavnyy, O. M., & Nykolaychuk, Ya. M. (2016). Theoretical Basis for Multidimensional Patterns Recognition in Hamming Space. *Scientific Bulletin of UNFU, 26(3), 361–367. Retrieved from:*

<http://nv.nltu.edu.ua/index.php/journal/article/view/225>

Nykolaichuk, Ya. M. (2012). *Kody polia Halua: teoriia ta zastosuvannia: monohrafiia*. Ternopil: Ternohraf, 575 p. [in Ukrainian].

Nykolaichuk, Ya. M. (2010). *Teoriia dzherel informatsii*. 2nd ed. Ternopil: TzOV "Ternohraf", 534 p. [in Ukrainian].

Nykolaichuk, Ya. M., & Zavediuk, T. O. (2010). Struktura ta funktsii rekurentnoho bioneirona dlia rozpoznavannia obraziv u Khemminhovomu prostori. *Zbirnyk naukovykh prats Buchatskoho instytutu menezhmentu i audytu, 6(2), 37–40*. Buchach: Hvyliia. [in Ukrainian].

Nykolaichuk, Ya. M., Vozna, N. Ya., & Pitukh, I. R. (2013). Structuring the movement of data in computer systems. Ternopil: Ternohraf, 284 p.

Vozna, N. Ya. (2014). Kontseptsiiia strukturyzatsii danykh v rozpodilenykh informatsiinykh systemakh. *Zbirnyk materialiv mizhnarodnoi koordynatsiinoi narady ISCM-2014 (pp. 55–57)*. Ternopil. [in Ukrainian].

Vozna, N. Ya., & Nykolaichuk, Ya. M. (2013). Osnovy teorii, funktsii ta zadachi strukturyzatsii danykh v informatsiinykh systemakh. *Pratsi mizhnarodnoi naukovoï konferentsii "Pytannia optymizatsii obchyslen (POO-XL)" (pp. 56–57)*. Kyiv: Instytut kibernetiky im. V. M. Hlushkova NAN Ukrainy. [in Ukrainian].

Vozna, N. Ya., Nykolaychuk, Ya. M., & Alishov, N. I. (2013). Conception and Theoretical Bases for Solving the Problem of Data Structuring in Information Computerized Systems. *Journal of Qafqaz University. Mathematics and Computer Science, 1(2), 112–117*. Baku: Azerbaijan.

Н. Я. Возна, А. И. Сыдор

Тернопольский национальный экономический университет, г. Тернополь, Украина

КРИТЕРИЙ СТРУКТУРНОЙ СЛОЖНОСТИ И ОЦЕНКИ РАСШИРЕННОГО ХЕММИНГОВОГО РАССТОЯНИЯ ДЛЯ ПРОСТРАНСТВЕННЫХ ОБЪЕКТОВ

Охарактеризовано современное состояние решения проблемы распознавания образов на основе теории Хеммингового расстояния. Изложены теоретические основы методов распознавания образов. Приведены недостатки известных методов распознавания образов в Хемминговом пространстве. Проанализированы методы оценки структурной сложности полифункциональных данных, предложены критерий и система атрибутов формализации структурированных данных. Приведены расчеты расширенной оценки Хеммингового расстояния для пространственных объектов. Приведена формализация известных методов распознавания одномерных образов. Приведены характеристики ансамблей для совокупностей сигналов оптимальных сигнальных пространств. Приведены решения задачи оценки структурной сложности полифункциональных данных на основе весовых коэффициентов оценок информативности. Представлена количественная, оптимальная оценка функционально-структурной сложности компонентов сложной системы в виде отношения. Реализованы методы распознавания образов на основе оценки структурной сложности. Предложены способы кодирования атрибутов и расчета Хеммингового расстояния на примере символика игральных карт. Приведены сравнительные диаграммы оценок Хеммингового расстояния для пар символов, используемых для различных методов нахождения Хеммингового расстояния. Охарактеризованы недостатки и преимущества различных предложенных способов кодирования атрибутов изображения.

Ключевые слова: сигнальное пространство; Евклидово расстояние; распознавание образов; информационные технологии; кодирование данных.

N. Ya. Vozna, A. I. Sydor

Ternopil National Economic University, Ternopil, Ukraine

CRITERION OF STRUCTURAL COMPLEXITY AND EVALUATION OF EXTENDED HEMMING DISTANCE FOR SPATIAL OBJECTS

Evaluation of extended Hamming distance that can be applied in various fields of knowledge is considered in the paper. Known methods of recognition in the Hamming space are not effective because of their using for images that are described by binary vectors, and do not address the applicability of different analytical expressions mutual correlation functions and capabilities of data encryption in various theoretical and numerical bases. In the course of the research the coefficient of structural complexity of components attributes of multifunctional data was determined; weight coefficients of informative estimation of their parameters were assigned as well. Structural complexity criteria of certain classes of images and methods for determining the Hamming distance based on difference distance of modular structural complexity units of images were developed. The analysis of the proposed methods was provided on the example of recognition of symbolism figures. The coding system symbols of playing cards are offered. The formulas for finding Hamming distance used in the proposed encoding systems were described and derived. Comparative diagrams of Hamming distance estimates for pairs of symbols used in the proposed methods of Hamming distance determination were given. The main advantages and disadvantages of each of the proposed methods for the Hamming distance determination were identified. The proposed de-

definition of Hamming distance encoding attributes of the objects, taking into account the complexity of the structural components, when determining of the extended Hamming distance qualitative features that most accurately and unambiguously characterize the object are considered that is a theoretical contribution to the theory of structural pattern recognition methods. Thus, the proposed methods of coding attributes and calculating of Hamming distance are the basis for using of such Hamming distance determining technologies for road signs, symbols chemicals and other two-dimensional images recognition.

Keywords: signal space; Euclidean distance; pattern recognition; Information Technology; data encryption.

Інформація про авторів:

Возна Наталя Ярославівна, канд. техн. наук, доцент, Тернопільський національний економічний університет, м. Тернопіль, Україна. **Email:** nvozna@ukr.net

Сидор Андрій Іванович, аспірант, Тернопільський національний економічний університет, м. Тернопіль, Україна. **Email:** Andrij9997@ukr.net