

ЗІСТАВЛЕННЯ З ШАБЛОНОМ НА БАЗІ ПОЛІНОМІВ КУНЧЕНКА В ДЕМОГРАФІЧНИХ ДАНИХ

В роботі досліджується можливість використання поліномів Кунченка для пошуку шаблонів в одновимірному сигналі з демографічними даними. Наведені результати статистичного експерименту, в якому метод на базі поліномів Кунченка порівнюється з класичними методами підрахунку взаємної кореляції та середньо квадратичного відхилення.

Ключові слова: зіставлення з шаблоном, поліноми Кунченка, демографічні дані.

В работе исследуется возможность применения полиномов Кунченко для поиска шаблонов в одномерном сигнале с демографическими данными. Приведены результаты статистического эксперимента, в котором метод на базе полиномов Кунченко сравнивается с классическими методами подсчета взаимной корреляции и среднеквадратического отклонения.

Ключевые слова: сопоставление с образцом, полиномы Кунченко, демографические данные.

This paper addresses using Kunchenko's polynomials as template matching method to recognize template in one-dimensional input signal with demographic data. Kunchenko's polynomials method is compared with classical methods – cross-correlation and sum of squared differences according to numerical statistical example.

Key words: template matching, Kunchenko's polynomials, vital statistics data.

Постановка проблеми

Класичною формою подання результатів будь-якого демографічного і, більше того, – статистичного спостереження є вихідні таблиці, в яких у боковик та підметі наведені певні атрибути (характеристики), а в матриці даних наведені кількісні значення респондентів, що задовольняють відповідним умовам. Найбільш відомою в демографії є таблиця статеві-вікового розкладу населення, яка в програмах перепису переважно більшості країн має перший номер. Як відзначав відомий польський демограф Едвард Россет (Edward Rosset), «віковий склад населення – це ніби живий пам'ятник усім демографічним фактам, що мали місце у минулі сто років» [1, с. 235]. Наприклад, табл. 1 містить статеві-віковий розподіл постійного населення України на 1 січня 2010 року [2].

Однак, коли ставиться задача пошуку певних закономірностей в розподілі наявних даних, то табличне представлення часто є надлишковим. Скажімо, якщо потрібно виявити існуючі періодичності чи інші структурні особливості статеві-вікового розподілу населення, то, принаймні, на перших етапах аналізу достатньо буде даних без їх ділення по статі. Іншими словами, замість всієї таблиці потрібно буде проаналізувати лише одну її колонку. Така ситуація характерна для більшості досліджень в статистиці та соціології. Впорядковуючи дані по часовим інтервалам, отримуємо ретроспективне представлення даних; взявши за основу КОАТУУ (класифікатор об'єктів адміністративно-територіального устрою України), будемо мати територіальний розподіл даних; по табл. 1 тривіально будуватиметься віковий розподіл тощо.

Таблиця 1

**Статеві-віковий розподіл постійного населення України на 1 січня 2010 року,
п'ятирічні вікові групи**

Вікові групи	Обидві статі	Чоловіки	Жінки
Всього	45782592	21107067	24675525
0-4	2357590	1213352	1144238
5-9	1961459	1008811	952648
10-14	2164511	1108446	1056065
15-19	2852303	1461446	1390857
20-24	3688815	1883496	1805319
25-29	3723728	1890806	1832922
30-34	3360706	1675102	1685604
35-39	3245340	1590932	1654408

40-44	3021268	1451098	1570170
45-49	3484781	1627904	1856877
50-54	3398490	1536111	1862379
55-59	3052164	1325128	1727036
60-64	2302575	948291	1354284
65-69	1971086	742016	1229070
70-74	2386099	846669	1539430
75-79	1271528	412075	859453
80-84	1040879	281420	759459
85-89	394550	83610	310940
90-94	81012	16076	64936
95-99	22525	3992	18533

Дані, отримані в результаті такого «розрізання» вихідної таблиці (проекування по одній з колонок), зручно представляти та аналізувати у вигляді графіку функції однієї змінної. Тому задачу пошуку певних закономірностей в розподілі демографічних даних часто можна звести до задачі виявлення заздалегідь визначених (з певних міркувань) образів в наявних одновимірних даних. Належний образ прийнято називати *шаблоном* чи *еталоном*. Відповідний напрям досліджень має назву *порівняння* чи *зіставлення з еталоном* (*template matching*) [3; 4]. Потреба мати справу з такого роду задачами постає в доволі різних сферах людської діяльності: від забезпечення машинного зору та розпізнавання людського почерку до інтелектуального аналізу даних.

Приміром, якщо взяти проекцію табл. 1 за колонкою «жінки», то на відповідному графіку, представленому на рис. 1, можна легко візуально виділити так зване покоління «бебі-бум» (baby-boom) – людей, котрі народилися приблизно з 1956 по 1965 рік. Для пошуку вікових діапазонів, схожих за динамікою зміни чисельності жінок в Україні, можна взяти в якості шаблону ділянку зазначеного графіку, показану на рис. 1 овалом. Цілком очікувано, що через 25 років після покоління «бебі-бум» маємо новий сплеск народжуваності. Але якщо потрібно здійснити аналогічний пошук по населенню окремих регіонів чи країн, або за людьми різного етнічного походження, то ця робота повинна бути автоматизована. Бо, наприклад, в США [5] чи в Японії [6] післявоєнний «бебі-бум» прийшовся (в п'ятирічних діапазонах) на період з 1945 по 1960 рік.

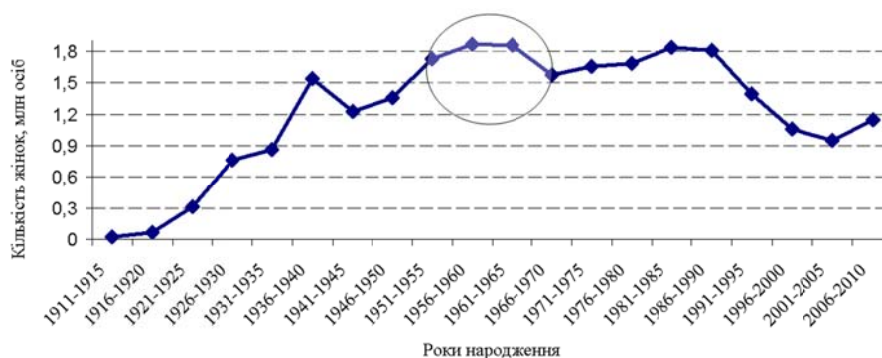


Рис. 1. Розподіл за віковими діапазонами кількості жінок, народжених в Україні

Аналіз останніх досліджень та публікацій

Існують два принципово різних підходи до організації пошуку окремих наперед визначених еталонів у деякому заданому масиві даних [7].

Перший підхід – це використання штучних нейронних мереж після їх навчання [8]. З математичної точки зору, навчання нейронної мережі є багатопараметричною задачею нелінійної оптимізації. Цей підхід вимагає обов'язкової наявності великої кількості прикладів з правильними відповідями, що є мало ймовірним під час оперативного дослідження демографічних даних.

Другий підхід полягає в порівнянні або певних особливостей еталону та вихідних даних (feature-based matching), або – безпосередньо всього еталону з частиною вихідних даних (template-based matching) [7]. Оскільки значну частину задач пошуку певних закономірностей в розподілі демографічних даних можна звести до аналізу відповідних графіків функції однієї змінної, то виділення особливостей в еталоні та вихідних даних (типу кутів чи країв зображення [9; 10]) не має особливого сенсу. Зупинимося на пошуку всього шаблону серед вихідної інформації.

Цей метод історично є одним із перших [11, с. 41]. Ще в 1932 р. радянським інженером В. Є. Агаповим була розроблена машина, що базувалася на цьому методі і призначалася для введення цифр до машино-рахувальної станції [12].

Постановка задачі зіставлення з еталоном включає в себе вихідні дані для аналізу, деякий шаблон для пошуку та, можливо, область, на якій потрібно здійснювати такий пошук. При цьому потрібно як правильно

обрати представлення вхідних даних, так і визначити критерії, за якими відбуватиметься їх зіставлення з шаблоном.

Демографічні дані з вихідної таблиці будемо представляти як одновимірний цифровий сигнал. За необхідності, можна розглядати й двовимірне графічне зображення, оскільки за останній час було запропоновано низку підходів зі зведення двовимірних зображень до одновимірних представлень (див., наприклад, [13]).

В нашому випадку результати класичної теорії статистичних рішень та теорії оцінювання [14-16] не можуть бути застосовані, бо відповідні методи потребують повного опису статистичних даних у вигляді щільності розподілу ймовірностей. В таких задачах, як розглядаємо, при зіставленні еталону з сигналом, зазвичай, використовують Евклідову відстань та метрики, побудовані на порівнянні сум (взаємна кореляція, сума абсолютних відхилень тощо) [17-20]. Але до поставленої задачі можна підійти й з іншої точки зору – спробувати наблизити початковий сигнал поліноміальною функцією, побудованою на основі еталону [21-23]. Тоді ділянки максимального наближення сигналу і побудованого поліному при перевищенні певного порогу можуть свідчити про наявність в сигналі заданого еталону. Даний підхід фактично базується на моментно-кумулянтному описі випадкових величин та процесів [24, 25].

Формулювання цілі статті

Загальновідомо, що поліном Тейлора є поліномом найкращого наближення функції в околі заданої точки. Але функція, що наближується, повинна мати в цьому околі похідні відповідного порядку. Значно слабші умови накладаються при розвиненні функції в ряд Фур'є по системі незалежних ортогональних функцій, що утворюють базис відповідного простору зі скалярним добутком. Проте з точки зору розпізнавання в вихідному сигналі якогось певного шаблону більш перспективним представляється використання так званих поліномів Кунченка [21; 23]. Пов'язано це з тим, що поліноми Кунченка будуються в особливому підпросторі Евклідова чи Гільбертова простора, який має породжуючий елемент [26]. В якості такого елемента можна взяти саме шуканий еталон. Тоді задача зіставлення з еталоном зведеться до більш простої задачі пошуку найближчого (в певній метриці) апроксимаційного поліному Кунченка [21].

Мета даної роботи полягає в здійсненні формалізації та експериментальної перевірки зазначеної вище схеми застосування поліномів Кунченка до пошуку шаблону в вихідних даних.

Побудова поліномів наближення в просторі з породжуючою функцією

Базуючись на роботах [26-28], введемо необхідні поняття.

Нехай задано так звану *породжуючу* функцію $f(x)$, визначену на проміжку $[a, d]$. Тоді, введемо впорядковану множину *породжених* функцій наступного вигляду:

$$u_v(x) = \varphi_v[f(x)], \quad x \in [a, d], \quad (1)$$

де $\varphi_v(\cdot)$ – дійсні функції, підібрані певним чином.

На цій множині очевидним чином можна ввести лінійні операції додавання та множення на число, тобто побудувати лінійний простір. Оскільки деякі з породжених функцій (1) можуть виявитися лінійно залежними, визначимо *просторово утворюючу множину* функцій, сформовану як об'єднання множини лінійно незалежних породжених функцій із (1). Відповідно, лінійний простір, утворений на такій множині, називається *лінійним простором над незалежними породженими функціями* (лінійним простором Кунченка) та позначається як *LFKi*.

Визначимо в просторі *LFKi* скалярний добуток двох елементів $u_v(x)$ і $u_k(x)$:

$$(u_v(x), u_k(x)) \equiv \int_a^d u_v(x) \cdot u_k(x) dx = \int_a^d \varphi_v[f(x)] \cdot \varphi_k[f(x)] dx. \quad (2)$$

Із формули (2) випливає визначення відстані ρ_{vk} між двома елементами $u_v(x)$ і $u_k(x)$ лінійного простору Кунченка як норми різниці двох функцій:

$$\rho_{vk}^2 \equiv \|u_v(x) - u_k(x)\|^2 = \int_a^d [u_v(x) - u_k(x)]^2 dx.$$

Якщо для всіх породжених функцій, що утворюють простір *LFKi*, можна ввести скалярний добуток (2), то такий простір називається *простором LFKi зі скалярним добутком*.

Простір *LFKi*, у якому множина породжених функцій складається *тільки* з лінійно незалежних, попарно (повністю чи частково) *неортогональних* функцій, називається *простором (Кунченка) з породжуючим елементом*. Під частковою неортогональністю породжених функцій розуміємо ситуацію, коли одна їхня непорожня частина є попарно неортогональною, а інша – попарно ортогональною.

Надалі, значення скалярного добутку (2) також називатимемо *корелятою* двох функцій $u_v(x)$ і $u_k(x)$ *розміру* (v, k) , яку позначатимемо наступним чином:

$$\Psi_{v,k} \equiv \int_a^d u_v(x) \cdot u_k(x) dx.$$

Також, виберемо деяку породжену функцію $u_b(x)$, яку назвемо *основною*. Тоді, можна утворити *поліном Кунченка r-го порядку* із породжених функцій $u_v(x) \neq b$, які назвемо *доповнювальними*:

$$P_r(x) \equiv \sum_{\substack{v=0 \\ v \neq b}}^r \alpha_v u_v(x) = \alpha_0 u_0(x) + \sum_{\substack{v=1 \\ v \neq b}}^r \alpha_v u_v(x). \quad (3)$$

Якщо розглянути задачу мінімізації відстані між поліномом Кунченка (3) та основною функцією $u_b(x)$:

$$\rho_{bP}^2 = \int_a^d [u_b(x) - P_r(x)]^2 dx, \quad (4)$$

то можна показати [26, с. 75–76], що коефіцієнти α_v , $v \neq 0$ знаходяться з наступної системи лінійних рівнянь:

$$\sum_{\substack{k=1 \\ k \neq b}}^r \alpha_k F_{v,k} = F_{v,b}, \quad v = \overline{1, r}, \quad v \neq b, \quad (5)$$

де *центровані кореляти* $F_{v,k} \equiv \Psi_{v,k} - \Psi_v \cdot \Psi_k \cdot \|u_0(x)\|^2$, а $\Psi_v \equiv \frac{\Psi_{v,0}}{\|u_0(x)\|^2} = \frac{\int_a^d u_v(x) \cdot u_0(x) dx}{\int_a^d u_0^2(x) dx}$.

Коефіцієнт α_0 повинен дорівнювати наступному виразу:

$$\alpha_0 = \Psi_b - \sum_{\substack{v=1 \\ v \neq b}}^r \alpha_v \Psi_v. \quad (6)$$

Якщо підставити коефіцієнти (5), (6) у вираз (4), то отримаємо:

$$\rho_{bP}^2 = F_{b,b} - J_r,$$

де так звана *інфоркуна* полінома (3) має вигляд $J_r \equiv \sum_{\substack{v=1 \\ v \neq b}}^r \alpha_v F_{v,b} = \sum_{\substack{v=1 \\ v \neq b}}^r \sum_{\substack{k=1 \\ k \neq b}}^r \alpha_v \alpha_k F_{v,k}$.

Нарешті, уведемо чисельний показник якості наближення поліномом (3) основної функції $u_b(x)$:

$$e_r = \frac{J_r}{\int_a^d (u_b(x) - \Psi_b \cdot u_0(x))^2 dx}. \quad (7)$$

Коефіцієнт (7) називається *коефіцієнтом ефективності* наближення полінома з доповнювальних функцій до основної функції. Що ближче даний коефіцієнт до 1, то кращим є наближення поліномом Кунченка відповідної функції [26, с. 86].

Варто зазначити, що оскільки множина доповнювальних функцій складається з повністю або частково неортогональних функцій, а основна функція не входить до множини доповнювальних, то ситуація, за якої коефіцієнт ефективності (7) дорівнює 1, можлива тільки в граничному випадку, коли порядок r полінома Кунченка прямує до нескінченності.

Ю. П. Кунченко та його учні застосовували розглянутий в даному пункті математичний апарат для вирішення задач, пов'язаних із статистичним оцінюванням параметрів випадкових сигналів [27; 28], для побудови систем статистичного розпізнавання, які базуються на властивості зменшення дисперсій випадкових величин при наближенні їх стохастичними функціональними поліномами [29], для поліноміального вирішення задачі розпізнавання сигналів на фоні негаусових завад [30] тощо. В даній роботі застосуємо апарат поліномів Кунченка для пошуку шаблонів в демографічних даних.

Застосування поліномів Кунченка для розв'язання задачі розпізнавання шаблону в сигналі

Розглянемо модифікацію описаного в попередньому пункті підходу до побудови поліномів Кунченка, виходячи з потреб задачі розпізнавання певного шаблону в сигналі [21–23, 31].

Спочатку визначимо, які вхідні дані потрібно мати для розв'язання поставленої задачі, а потім укажемо для них місце в схемі, запропонованій Ю. П. Кунченком [26].

Отже, для виконання розпізнавання конкретного шаблону в деякому сигналі слід мати сам *сигнал* $s(x)$, визначений на проміжку $[a, d]$, а також певний функціональний *шаблон* $\psi(x)$, $x \in [a, d]$. За таких початкових даних нам потрібно отримати відповідну кількісну оцінку схожості обраного шаблону на сигнал для аналізу.

Тому [31, с. 72],

- в якості *основної функції* візьмемо початковий *сигнал*: $u_b(x) = s(x)$;
- в якості *доповнювальних функцій* потрібно взяти лінійно незалежні та частково неортогональні на проміжку $[a, d]$ функціональні перетворення *шаблону*: $u_v(x) = \varphi_v[\psi(x)]$;
- кількісною характеристикою схожості шаблону та сигналу слугуватиме коефіцієнт ефективності (7).

Проте, дану схему можна застосувати до оцінки схожості двох функцій, визначених тільки на одному й тому ж проміжку. У багатьох практичних задачах довжина шаблону значно коротша за довжину сигналу, тому існує потреба у визначенні ділянок сигналу, які відповідають якнайбільшій близькості до шаблону.

Розглянемо сигнал $s(x)$, який визначено на проміжку $[a, d]$, та шаблон $\psi(x)$, визначений на іншому проміжку $[i, j] \subseteq [a, d]$. При цьому для нас не є принциповим, яких конкретно значень набувають числа i та j ; важливо лише, що обидва ці числа лежать на проміжку $[a, d]$, і довжина проміжку $[i, j]$ фіксована. Тоді, можна здійснити процедуру побудови полінома Кунченка для певного вікна, тобто ділянки початкового сигналу $s(x)$ довжиною $|i - j|$.

Таким чином, отримаємо множину значень коефіцієнтів ефективності:

$$E = \{e(x)\} \quad \forall x \in [a, d - j + i],$$

яку назовемо *ефективною множиною*, а графік, що представляє значення з цієї множини, – *ефектограмою*. Аналізуючи ефектограму, можна визначити, які ділянки сигналу і наскільки відповідають запропонованому шаблону [31, с. 72].

Опис модельного прикладу

Проведемо порівняння запропонованого методу з іншими зіставними методами класу template matching на модельному прикладі.

Для порівняння візьмемо два класичні методи зіставлення еталону з сигналом [17-20], котрі використовують Евклідову відстань та полягають в пошуку або екстремумів взаємної кореляції (кореляційний метод), або мінімумів суми квадратів відхилень (метод найменших квадратів) шаблону та сигналу.

Модельний приклад побудуємо наступним чином. В якості шаблону візьмемо класичну гаусіану $\psi_1(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$. При цьому, через те, що нас цікавлять у першу чергу локальні особливості відповідного сигналу, обмежимо носій експоненти відрізком $[-3, 3]$. Відліки з цього відрізка візьмемо з кроком 0,01. Графічно відповідний шаблон представлено на рис. 2.

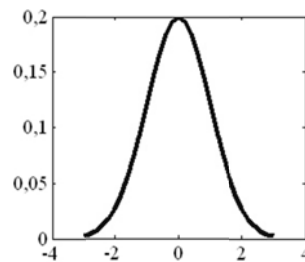


Рис. 2. Графік шаблону для модельного експерименту

Вихідний сигнал $s(x)$ будемо генерувати як суму 10 гаусіан:

$$s(x) = \sum_{i=1}^{10} \frac{1}{\sqrt{2\pi}\sigma_i} e^{-(x-\mu_i)^2/(2\sigma_i^2)},$$

де $\mu_1 = 0$, $\mu_k = \mu_{k-1} + 4(\sigma_{k-1} + \sigma_k)$, $k = \overline{2, 10}$, $\sigma_i = \text{rand}(0, 2; 2)$, $i = \overline{1, 10}$, $\text{rand}(0, 2; 2)$ – функція, що повертає псевдовипадкове число, рівномірно розподілене на інтервалі $(0, 2; 2)$.

Приклад вихідного сигналу наведено на рис. 3. Для розрахунку точок гаусіан використовувалася вбудована функція gaussmf системи комп'ютерної математики MATLAB. Область визначення вихідного сигналу, також як і шаблону, була дискретизована відліками з кроком 0,01.

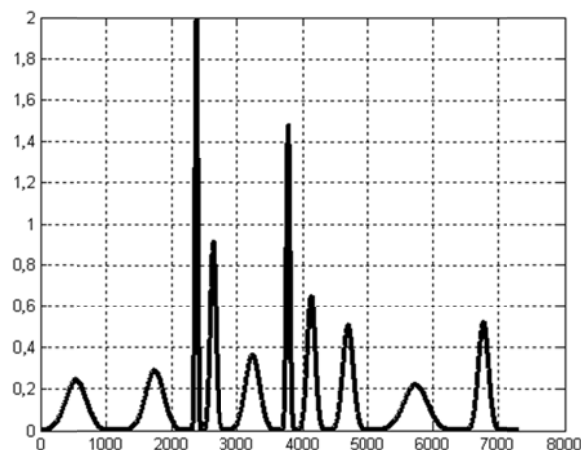


Рис. 3. Приклад вихідного сигналу із 10 гаусіан

Для більшої реалістичності до вихідного сигналу було додано білий шум, причому відношення сигнал/шум було взяте рівним 30 дБ. Приклад такого сигналу, побудованого по сигналу, наведеному на рис. 3, представлено на рис. 4.

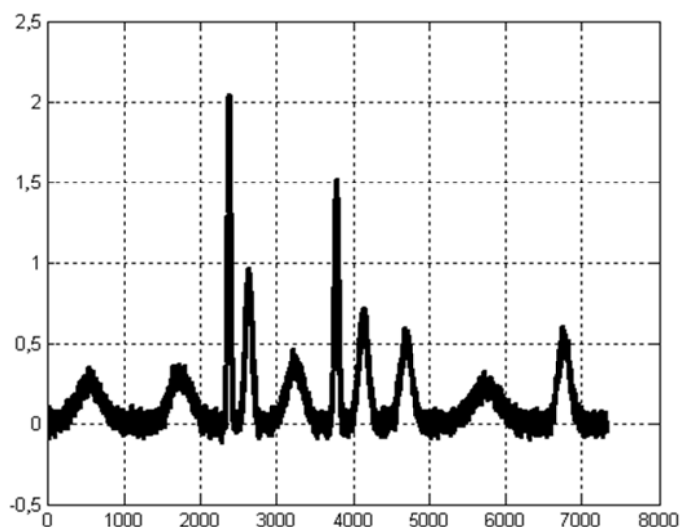


Рис. 4. Приклад вихідного сигналу із 10 гаусіан з додаванням білого шуму

Було проведено 1000 описаних експериментів, результати яких зведено на рис. 5 і 6.

Рис. 5 містить відсоток виявлених у вхідному сигналі шаблонів в залежності від величини порогу для кожного з трьох порівнюємих методів: метода на базі поліномів Кунченка, кореляційного метода та метода пошуку мінімальної суми квадратів відхилень. В якості характеристик, перевищення значення котрих над порогом рахувалося як знайдення шаблону, відповідно бралися: коефіцієнт ефективності, розрахований згідно (7), коефіцієнт взаємної кореляції Пірсона [19] та нормована сума квадратів відхилень між значеннями сигналу в кожному відліку вікна та значеннями шаблону.

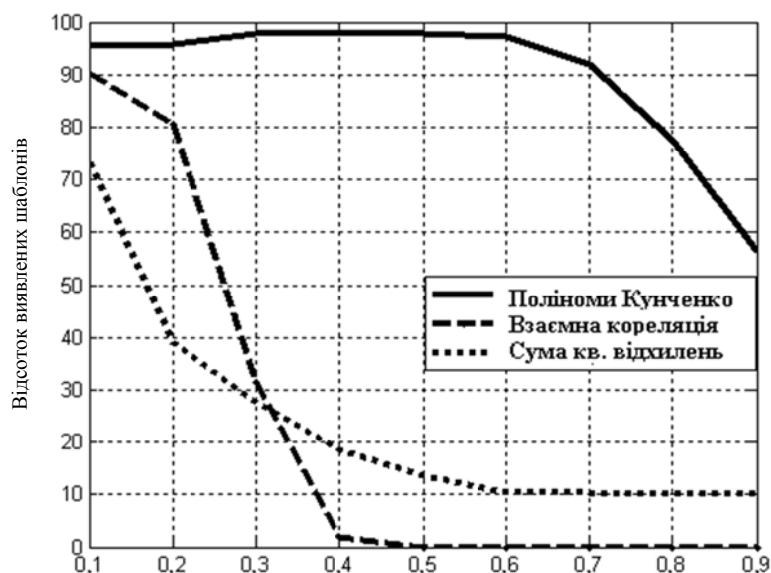


Рис. 5. Відсоток виявлених шаблонів в залежності від величини порогу

Рис. 6 містить побудований за тим же самим принципом, що й на рис. 5, відсоток хибних спрацьовувань зазначених методів під час пошуку шаблонів в сигналі, причому за 100 % бралася кількість усіх шаблонів під час проведення всіх 1000 експериментів.

Отримані результати свідчать, що за оптимального підбору значення порогу метод, який базується на побудові поліномів наближення в просторі з породжуючим елементом, дозволяє в середньому виявити не менше 98 % шаблонів, присутніх в вихідному сигналі, якщо відношення сигнал/шум дорівнює 30 дБ. Величина відсотку хибних спрацьовувань, тобто кількості ділянок сигналу, які помилково були визначені як

близькі до шаблону, відносно загальної кількості шаблонів в сигналі, при порозі, що дорівнює 0,6, склала приблизно 3 %.

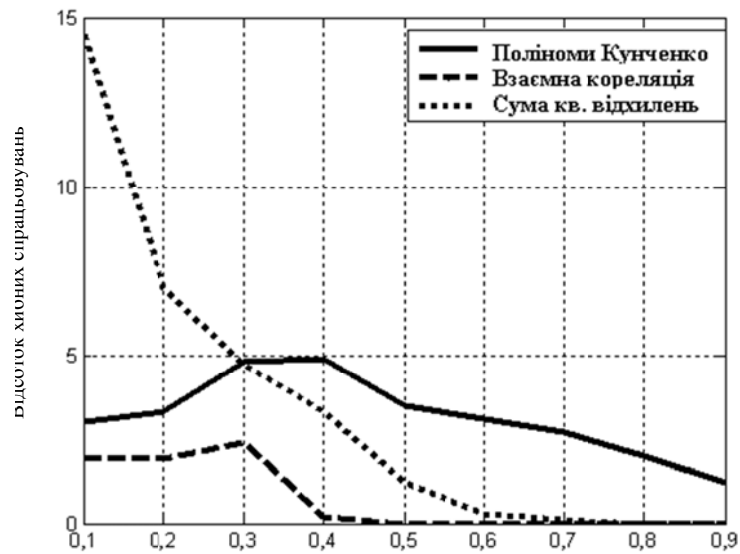


Рис. 6. Відсоток хибних спрацьовувань під час пошуку шаблонів в сигналі

Близький результат дає й класичний метод, що базується на підрахунку взаємної кореляції шаблону і сигналу (див. рис. 5 і 6).

Висновки

Статистичний експеримент показав, що метод зіставлення з еталоном, побудований на основі поліномів Кунченка, дає кращі результати розпізнавання не масштабованого шаблону і дозволяє знаходити шаблони з більшою ймовірністю, ніж класичні методи, що базуються на підрахунку взаємної кореляції чи сумі квадратів відхилення шаблону від сигналу. При виборі найкращого порогу запропонований метод дає майже стільки ж хибних спрацьовувань, як і класичні методи.

В ряді досліджень, наприклад, демографічних, суттєву роль відіграє можливість не просто виявлення в сигналі вихідних даних загальних закономірностей, характерних для певного еталону (опуклість чи вгнутість, фіксована зміна монотонності тощо), а, насамперед, – знаходження ділянок сигналу, що мають ідентичну чи близьку до еталону форму. Метод на базі поліномів Кунченка, тобто поліномів наближення в просторі з породжуючою функцією, потенційно повинен давати гарні результати при розв'язанні таких задач, якщо еталон взяти в якості породжуючої функції відповідного простору. Детальний аналіз факторів, що впливають на якість розпізнавання за допомогою поліномів Кунченка, є предметом подальших досліджень.

ЛИТЕРАТУРА

1. Россет Э. Процесс старения населения : [пер. с англ.] / Э. Россет. – М. : Статистика, 1968. – 510 с.
2. Таблица 0204. Розподіл постійного населення за статтю, віковими групами та типом поселень (УО) [Електронний ресурс] / Державний комітет статистики України. – Режим доступу : http://stat6.stat.lviv.ua/ukrcensus/Dialog/varval.asp?ma=000_0204&ti=0204. Розподіл постійного населення за статтю, віковими групами та типом поселень (УО)&path=../Database/Population/02/02/&lang=1
3. Ballard D. H. Computer Vision / D. H. Ballard, C. M. Brown. – New Jersey : Prentice-Hall, Englewood Cliffs, 1982. – 527 p.
4. Brunelli R. Template Matching Techniques in Computer Vision: Theory and Practice / Roberto Brunelli. – Chippenham : John Wiley & Sons, 2009. – 348 p.
5. Strauss W. Generations: The history of America's future, 1584 to 2069 / W. Strauss, N. Howe. – N. Y. : Harper Perennial, 1992. – 538 p.
6. Nitta H. Capitalizing on Retirement of Japan's First Baby-Boomers [Electronic resource] / H. Nitta // JETRO Japan Economic Report, April-May, 2006 / Japan External Trade Organization. – Mode of access : http://www.jetro.go.jp/en/reports/market/pdf/2006_18_u.
7. Fellenz W. A. Comparing Template-based, Feature-based and Supervised Classification of Facial Expressions from Static Images / W. A. Fellenz, J. G. Taylor, N. Tsapatsoulis, S. Kollias // The 3rd IMACS, IEEE, World SES, Multiconference on Circuits, Systems, Communications and Computers (CSCC'99), July 4–8, 1999 : proceedings. – Athens : World Scientific Engineering Society Press, 1999. – P. 5331–5336.
8. Хайкин С. Нейронные сети: полный курс : [пер. с англ.] / Саймон Хайкин. – [2-е изд.]. – М. : ИД «Вильямс», 2011. – 1104 с.
9. Margalit A. Using feature probabilities to reduce the expected computational cost of template matching / A. Margalit, A. Rosenfeld // Computer Vision, Graphics and Image Processing. – 1990. – № 52. – P. 110–123.
10. Fredriksson K. Sequential and indexed two-dimensional combinatorial template matching allowing rotations / K. Fredriksson, G. Navarro // Theoretical Computer Science. – 2005. – № 347. – P. 239–275.
11. Васильев В. И. Распознающие системы : справочник / В. И. Васильев. – [2-е изд.]. – К. : Наукова думка, 1983. – 422 с.
12. Саплин М. С. Фотоэлектрическое устройство, воспринимающее цифровые печатные знаки / М. С. Саплин // Электрон. вычисл. машины. – 1960. – № 1. – С. 110–123.
13. Lin Y.-H. Template matching using the parametric template vector with translation, rotation and scale invariance / Y.-H. Lin, C.-H. Chen // Pattern Recognition. – 2008. – Vol. 41, № 7. – P. 2413–2421.

14. Фукунага К. Введение в статистическую теорию распознавания образов : [пер. с англ.] / К. Фукунага. – М. : Наука, 1979. – 368 с.
15. Левин Б. Р. Теоретические основы статистической радиотехники / Б. Р. Левин. – М. : Сов. Радио, 1969. – 752 с.
16. Трифонов С. И. Совместное различение сигналов и оценка их параметров / С. И. Трифонов, А. Н. Шинаков. – СПб : Политехника, 1996. – 425 с.
17. Пратт У. Цифровая обработка изображений : [пер. с англ.] / У. Пратт. – М. : Мир, 1982. – Т. 1. – 310 с.
18. Rosenfeld A. Digital Picture Processing / A. Rosenfeld, A. C. Kak. – [2nd ed.]. – N. Y. : Academic Press, 1976. – 469 p. – (Computer Science and Applied Mathematics series).
19. Tsai D.-M. Fast normalized cross correlation for defect detection / D.-M. Tsai, C.-T. Lin // Pattern Recognition Letters. – 2003. – Vol. 24, Issue 15. – P. 2625–2631.
20. Lewis J. P. Fast Template Matching / J. P. Lewis // Vision Interface. – 1995. – Vol. 95. – P. 120–123.
21. Чертов О. Р. Поліноми Кунченка для розпізнавання образів / О. Р. Чертов // Вісник НТУУ «КПІ». Інформатика, управління та обчислювальна техніка : зб. наук. пр. – К. : Век+, 2009. – № 51. – С. 105–110.
22. Group methods of data processing : monograph / O. Chertov, D. Tavrov, D. Pavlov [et al.] ; [ed. O. Chertov]. – Raleigh : Lulu.com, 2010. – 155 p.
23. Чертов О. Р. Застосування поліномів Кунченка при моделюванні складних систем / О. Р. Чертов, Д. Г. Павлов, М. В. Александрова // Вісник Херсонського національного технічного університету. – 2010. – № 3 (39). – С. 503–508.
24. Малахов А. Н. Кумулянтный анализ негауссовских случайных процессов и их преобразований / А. Н. Малахов. – М. : Сов. радио, 1978. – 376 с.
25. Hu M. K. Visual pattern recognition by moment invariants / M. K. Hu // IEEE Transactions on Information Theory. – 1962. – № 8. – P. 179–187.
26. Кунченко Ю. П. Приближения в пространстве с порождающим элементом / Ю. П. Кунченко. – К. : Наук. думка, 2003. – 243 с.
27. Кунченко Ю. П. Стохастические полиномы / Ю. П. Кунченко. – К. : Наук. думка, 2006. – 275 с.
28. Кунченко Ю. П. Оценка параметров случайных величин методом максимизации полинома / Ю. П. Кунченко, Ю. Г. Лега – К. : Наук. думка, 1992. – 180 с.
29. Заболотний С. В. Статистичне розпізнавання образів на основі розкладу в просторі з порідним елементом // Вісник НУ Львівська політехніка «Комп'ютерні науки та інформаційні технології». – 2009. – № 638. – С. 118–123.
30. Палагін В. В. Поліноміальне вирішення задач розпізнавання випадкових сигналів / В. В. Палагін, О. М. Жила // Вісник Черкаського державного технологічного університету. – 2008. – № 2. – С. 31–35.
31. Чертов О. Р. Застосування поліномів Кунченка для аналізу статистичних даних / О. Р. Чертов, Д. Ю. Тавров // Східноєвропейський журнал передових технологій. – 2010. – № 4/4 (46). – С. 70–75.