

Харківський національний університет радіоелектроніки

В.О. Гороховатський, С.В. Гадецька

**Статистичне оброблення та аналіз даних у
структурних методах класифікації зображень**

Монографія

Харків 2020

УДК 004.932.2:004.93'1

Гороховатський В.О., Гадецька С.В. Статистичне оброблення та аналіз даних у структурних методах класифікації зображень (монографія). – Харків: ФОП Панов А.Н., 2020. – 128 с.
ISBN 978-617-7859-69-6

Монографія містить матеріали досліджень у сфері проблем комп'ютерного зору задля створення інтелектуальних систем розпізнавання зображень на основі використання структурних методів. Основна увага приділяється впровадженню апарату статистичного аналізу даних, що є фундаментальною основою функціонування сучасних систем розпізнавання. Запропоновано новітні методи класифікації зображень на підґрунті опису у вигляді множини дескрипторів ключових точок, проведено їх статистичне обґрунтування. Впроваджено ефективні підходи просторового аналізу даних описів із використанням фрагментного оброблення. Розроблено ряд нових перспективних просторів даних для систем розпізнавання, де досягається суттєве поліпшення показників швидкодії та достовірності, вивчені статистичні підходи до оптимізації параметрів запропонованих методів. Детально описано результати експериментальних досліджень і застосування розроблених методів при вирішенні задач класифікації для прикладних баз зображень.

Для науковців, викладачів, аспірантів, студентів та всіх читачів, хто цікавиться проблемами створення новітніх інформаційних технологій та інтелектуальних методів аналізу чи оброблення даних, а також їх впровадженням у системах комп'ютерного зору.

УДК 004.932.2:004.93'1

ISBN 978-617-7859-69-6

DOI: 10.30837/978-617-7859-69-6

© Гороховатський В.О., Гадецька С.В., 2020

Зміст

Вступ.....	5
1. Структурні методи класифікації зображень за описом множини дескрипторів ключових точок.....	10
1.1 Формалізація структурного методу класифікації.....	10
1.2 Аналіз і вибір детекторів КТ.....	14
1.3 Грануляція значень множини дескрипторів	17
Висновки.....	28
2. Статистичний аналіз даних для описів у вигляді системи кластерів.....	29
2.1 Класифікація на засадах баєсовської теорії прийняття рішень.....	29
2.2 Класифікація на підставі статистичних мір релевантності.....	40
2.2.1 Баєсовський класифікатор.....	41
2.2.2 Метричні класифікатори.....	43
2.2.3 Порівняльний аналіз прикладних властивостей мір релевантності.....	46
Висновки.....	51
3. Визначення релевантності зображень шляхом застосування статистичного аналізу для фрагментного подання їх описів.....	52
3.1 Формалізація побудови системи статистичних розподілів.....	52
3.2 Оцінювання значущості прийняття класифікаційних рішень для моделі блочного подання.....	54
3.3 Обчислення релевантності та класифікація за значеннями розподілів.....	63
3.4 Просторово-статистичний аналіз та оброблення ансамблю фрагментів...64	
Висновки.....	76
4. Результати комп'ютерного моделювання розроблених методів.....	77
4.1. Експериментальний аналіз зіставлення розподілів фрагментів.....	77
4.2. Розрахункова оцінка ступеня відмінності експериментальних даних.....	80
4.3 Результати моделювання засобів просторового оброблення.....	82
4.4 Аналіз комп'ютерного моделювання ієрархічних ознак.....	85

Висновки.....	88
5. Прикладні системи статистичних ознак для структурних описів.....	90
5.1 Побудова схеми зважування tf-idf.....	91
5.2 Лінгвістичний аналіз структурних описів.....	93
5.3 Застосування медіани для трансформації і зіставлення описів.....	95
5.4 Аналіз результатів комп'ютерного моделювання.....	99
5.5 Обговорення досліджень модифікацій методів.....	105
Висновки.....	115
Список літературних джерел.....	117

Вступ

У сучасних системах комп'ютерного зору набули популярності та практичного застосування структурні методи розпізнавання візуальних об'єктів, що засновані на локальних ознаках зображення. Такі методи базуються на визначенні множини ключових точок (КТ) та їх описові у вигляді числового чи бінарного вектора – дескриптора, що відображає властивості вмісту яскравостей зображення для локальних околиць КТ [1-10]. Такий підхід дає можливість комп'ютерній програмі працювати із візуальними образами подібно людині, зір якої теж часто ґрунтується на особливих локальних прикметах зображення.

Значення дескриптора, як правило, є інваріантним стосовно групи геометричних перетворень об'єктів на зображенні (зміщення, поворот, масштабування), а кількість утворених дескрипторів, що формують опис, повинна бути достатньою для прийняття результативного рішення відносно розрізнення розпізнаваних об'єктів [11-20]. Розпізнавання (класифікація) при застосуванні структурних методів полягає в обчисленні міри релевантності між описами об'єкта та еталона, образи яких задані у вигляді множин векторів-дескрипторів, та оптимізації значень цієї міри для бази еталонів.

Подальший розвиток нинішніх систем комп'ютерного зору безпосередньо пов'язаний зі створенням все більш ефективних систем ознак та моделей представлення зображень задля синтезу новітніх методів класифікації. Основою для розпізнавання зображень візуальних об'єктів, описаних множинами КТ, є метричні відношення, що відображають схожість як окремих КТ, так і описів візуальних об'єктів загалом. Це фактично встановлює деяку топологію для представлення зображень чи їх фрагментів у побудованому просторі ознак.

Відповідно до структурного підходу, зображення об'єктів, які складають зміст зображення сцени, складають образи об'єктів із окремих КТ

як своєрідних цеглинок. Переміщення сформованих КТ у просторі координат адекватно відображає фізичне пересування об'єкта. Топологія множини КТ також певною мірою відображає фізичну топологію пізнаваних об'єктів у побудованому просторі образів.

Перспективною ідеєю в аспекті скорочення обсягу обчислювальних витрат для процесу класифікації візуальних об'єктів, які містить аналізоване зображення, є побудова деякої множини опорних точок у просторі дескрипторів КТ, які формують будь-який структурний опис [2, 10, 11, 21, 22]. Комбінацією числа КТ, близьких до опорних точок у просторі їх значень (дескрипторів КТ), можна універсально отримати опис довільного образу об'єкта. Схожі можливості також забезпечує апарат мультимножин [34, 35].

Розвиненням ідеї кластерного уявлення для структурних описів у межах прикладної бази зображень є безпосередня побудова правила класифікації, яке за значенням дескриптора буде відносити КТ до фіксованого еталонного класу. У результаті аналізу накопичується певна кількість голосів КТ, за якою згодом визначається номер класу для аналізованого об'єкта. Для близьких за змістом зображень таке правило із високою ефективністю важко реалізувати, тому що схожі зображення часто містять і чимало однакових дескрипторів опису. Однак, наші дослідження показують, що для достатнього різноманіття баз зображень синтез класифікаційного правила щодо окремих дескрипторів КТ все-таки може дати непогані прикладні результати [22-25]. Викликає також науковий інтерес застосування інтелектуальних технологій навчання у просторі описів як множин КТ для досягнення значимих показників класифікації [21, 56].

Як правило, по своїй природі оброблювані дані у комп'ютерному зорі мають багатовимірний ймовірнісний розподіл. Але іноді їх можна подати у вигляді системи розподілів одновимірних величин, що значно спрощує класифікацію та суттєво знижує необхідні обчислювальні затрати [12].

Аналіз багатовимірних сигналів шляхом представлення їх як множини фрагментів досить популярний у методах оброблення зображень [26-30], так як часто найважливіша інформація про образи візуальних об'єктів зосереджується у окремих деталях.

Бітова природа дескрипторів КТ у просторі бінарних векторів дає можливість впровадити подання та аналіз окремого дескриптора як ланцюжка елементів (наприклад, байтів чи набору бітів), діапазон значень яких відомий. Це дає змогу здійснювати статистичний аналіз даних з урахуванням внутрішнього змісту наявного дескрипторного опису об'єкту. З точки зору важливості інформації всі елементи ланцюжка рівноцінні, але місце їх розміщення у складі дескриптора фіксоване, тому є можливість аналізувати чи обробляти упорядковані послідовності елементів. Ланцюжкова структура допускає застосування статистичних підходів інтелектуального аналізу, заснованих на ймовірнісних оцінках наявних значень даних, щоб прийняти рішення про віднесення об'єкту з описом у вигляді множини дескрипторів до відповідного класу [4. 7-10, 12-18].

Статистичне подання є одним із найбільш популярних інструментів у сучасному інтелектуальному аналізі даних задля виявлення закономірностей чи встановлення системи знань, що містять дані [12-18]. Будь-які процедури голосування структурних ознак зображень базуються на статистичному аналізі даних [21, 28]. Методи статистичного аналізу ґрунтуються на виявленні у змісті аналізованої інформації суттєво значимих характеристик, вартість яких оцінюють значенням параметру частоти зустрічальності як оцінки ймовірності аналізованої події чи закономірності. Процедури навчання та самонавчання в інтелектуальних системах в основному теж базуються на визначенні найбільш вживаних компонентів даних [19, 21]. На підставі аналізу скінченного об'єму наявних даних здійснюють апроксимацію ймовірнісного розподілу, що дає змогу розпізнавати їх образи. Важливим аспектом, що досить точно відображає узагальнені властивості чи

характеристики даних, є безпосереднє використання значень розподілів у методах розпізнавання [18, 20].

Один із можливих підходів побудови класифікатора засновано на використанні таких статистичних характеристик, як математичне очікування, дисперсія, оцінки медіани та ін. [24, 25]. Але більш інформативним є все-таки використання безпосередньо самих значень розподілів, як це впроваджено у методах інтелектуального аналізу. Це дає змогу чуттєвіше врахувати відмінності та особливості значень даних, що відображають розрізнявальні властивості розпізнаваних класів для зображень.

Беззаперечно, що основою сучасного апарату data science є статистичні методи, які теоретично та практично розвиваються вже протягом довгого часу [50-85]. Їх перевагами є строге теоретичне обґрунтування та наявність розроблених програмних середовищ для практичного впровадження. Одним із найбільш застосованих у задачах комп'ютерного зору є програмні засоби бібліотеки OpenCV [86]. Бібліотека OpenCV містить понад 2500 оптимізованих класичних та сучасних алгоритмів аналізу зображень і машинного навчання. Можливості таких сучасних програмних бібліотек комп'ютерного зору, як Open CV, забезпечують вирішення ряду нагальних практичних задач: аналіз вмісту зображень, пошук та розпізнавання заданих об'єктів, виявлення тексту, відстеження рухів об'єктів, виявлення спільних елементів на порівнюваних зображеннях, реалізація методів навчання для прикладних баз відеоданих тощо.

Матеріал монографії в основному ґрунтується на недавніх дослідженнях авторів [1-25, 31-36, 56, 61-63], де детально викладене теоретичне підґрунтя та особливості прикладного застосування статистичного апарату стосовно задачі структурної класифікації зображень. Проведені дослідження включають порівняльний аналіз розроблених методів із традиційними метричними підходами [26-30] та базовими структурними

методами в аспекті експериментально підтверджених показників результативності та швидкодії [41-45].

Автори висловлюють щире подяку аспірантам та студентам Харківського національного університету радіоелектроніки, які є співавторами ряду публікацій і на протязі тривалих років приймали дієву участь у розробленні програмних продуктів, здійсненні програмного моделювання та дослідженні результативності розроблених методів з впровадженням у прикладних базах зображень: Берестовському А.Є., Кузьміну С.В., Кулікову Ю.О., Поляковій Т.В., Пономаренку Р.П., Пупченку Д.В., Сірик Т.О., Солодченку К.Г., Столярову В.С. та іншим.

Монографію автори присвячують пам'яті відомого вченого, Заслуженого діяча науки і техніки України, доктора технічних наук, професора Путятіна Євгенія Петровича, який нещодавно пішов із життя. Професор Путятін Є.П. створив власну наукову школу, на протязі 1977-2020 рр. був завідувачем кафедри інформатики Харківського національного університету радіоелектроніки, дослідження монографії належать до здобутків наукової школи професора Путятіна Є.П.

1. Структурні методи класифікації зображень за описом множини дескрипторів ключових точок

1.1. Формалізація структурного методу класифікації

Дотримуючись традиційного підходу [28], який актуальний і в даний час, сформулюємо математичну постановку задачі класифікації візуального об'єкта, представленого описом у просторі дескрипторів КТ [5]. Маємо два стрижневих простори даних: простір сигналів, в якому формуються розпізнавані образи – зображення об'єктів, і простір ознак, де реалізується прийняття рішення про клас образів. Класи об'єктів задані у просторі сигналів, а визначаються і розрізняються у просторі ознак. На рис. 1.1 наведено приклад зображення і візуальне відображення координат КТ, сформованих детектором.



Рисунок 1.1. Зображення та його множина координат КТ

Нехай Ω – деякий простір образів (зображень); $\omega \in \Omega$ – образ (візуальний об'єкт, зображення); $M[J] = \{1, 2, \dots, J\}$ – множина номерів для класів образів $\Omega_1, \Omega_2, \dots, \Omega_J$ таких, що $\Omega = \cup_{i=1}^J \Omega_i$, $\Omega_i \cap \Omega_k = \emptyset$, \emptyset – пуста

множина; $g : \Omega \rightarrow M$ – невідома індикаторна функція, яка ставить у відповідність кожному образу $\omega \in \Omega$ деякий номер $j \in M$ його класу.

Розглянемо тепер $Z = \{z \mid z \in R^n\}$ – простір описів образів, де опис – це скінченна множина числових векторів (дескрипторів КТ). задамо відображення $\Theta : \Omega \rightarrow Z$ із простору Ω у простір Z . Кожному $\omega \in \Omega$ відображення Θ ставить у відповідність його опис $z(\omega) = \{z_i\}_{i=1}^s$. Опис $z(\omega)$ у просторі Z – це скінченна множина: $z(\omega) = \{z_i\}_{i=1}^s$, $s = \text{card } z(\omega)$ – потужність $z(\omega)$; Z_1, Z_2, \dots, Z_J – підмножини у просторі Z , такі, що $Z_i \cap Z_k = \emptyset$ і $Z = \cup_{i=1}^J Z_i$. Передбачається, що ці підмножини отримані дією відображення $\Theta : \Omega_k \rightarrow Z_k$.

Тепер розглянемо $\hat{g} : Z \rightarrow M$ – вирішальне правило, яке множині $z(\omega)$ ознак образу ставить у відповідність номер класу образу. Процес класифікації – це ланцюжок перетворень даних із простору образів у множину класів, має вид

$$\Omega \rightarrow \Theta(\Omega) \rightarrow Z \rightarrow \hat{g}(Z) \rightarrow M. \quad (1.1)$$

Головною задачею процесу (1.1) є побудова правила \hat{g} , базуючись на поданні у просторі Z , отриманому застосуванням Θ . Вибір Z та $z(\omega)$ прямо впливає на ефективність правила \hat{g} . При цьому критерієм якості класифікації виступає результативність (правильність) прийняття рішення щодо елементів вихідного простору образів Ω .

У задачі розпізнавання, де описи представлені сукупністями КТ зображення, множини описів $z(\omega)$ формуються детекторами для побудови КТ: SIFT, SURF, ORB, BRISK и др. [37-40]. Конкретно для метода SURF $z(\omega) = \{z_i \mid z_i \in R^{64}\}_{i=1}^s$. В окремому випадку множина $z(\omega)$ подається у вигляді єдиного вектору, тоді Z трактується як простір векторів.

Обмеження традиційної постановки не враховують, що образи у просторі Ω , як і їх описи у просторі Z , можуть мати непорожній перетин. Цей нетривіальний аналіз забезпечують сучасні методи, засновані, зокрема, на імовірнісному або нечіткому уявленнях [40].

На попередньому етапі здійснюємо формування бази еталонних образів і побудову для кожного з них структурних описів Z_1, Z_2, \dots, Z_J .

Базовий метод класифікації полягає в наступному [1, 5].

1. Побудова опису $O \in Z$ для распізнаваного об'єкту.
2. Обчислення значень релевантності $d_j(Z_j, O)$ для опису O на множині Z_1, Z_2, \dots, Z_J описів еталонів.

3. Оптимізація на множині значень $d_j(Z_j, O)$, $j = 1, \dots, J$, і визначення номеру ν класу еталона $\nu = \arg \operatorname{opt}_M d_j(Z_j, O)$ з найкращим значенням релевантності (максимум для подібності або мінімум для відстані).

Побудуємо класифікаційне правило $\hat{g}: Z \rightarrow M$ шляхом формування центрів класів $\{m_j\}_{j=1}^J$, $m_j \in Z$ на навчальній множині $Z(J) = \{Z_1, Z_2, \dots, Z_J\}$. Квантування за допомогою множини опорних точок визначає дискретну модель простору ознак. Класифікацію будь-якого $x \in Z$ здійснюватимемо конкурентним способом

$$\nu = \arg \min_{j=1, \dots, J} \rho(x, m_j), \quad (1.2)$$

де ν – визначений клас КТ, $\rho(x, m_j)$ – відстань між дескрипторами КТ.

Зрозуміло, що при распізнаванні на Z правило (1.2) на практиці потрібно доповнити логічним аналізом із застосуванням параметра порогу [5-9]: віднесення дескриптора КТ до класу ν здійснюється тільки у ситуації, якщо для вектора $x \in Z$ буде також виконана умова $\rho(x, m_j) \leq \delta$, де δ – заданий поріг еквівалентності елементів. При відсутності такого логічного

аналізу класифікованим може бути зовсім не релевантний елемент, хоча мінімум в (1.2) завжди буде досягнуто.

У результаті застосування (1.2) до фіксованого опису $z(\omega)$ отримаємо трансформований опис як вектор $h[z(\omega)]=(h_1, h_2, \dots, h_J)$, де h_j – цілі числа. Цей вектор відповідає розподілу елементів вхідної множини дескрипторів за еталонними класами. Аналогічні уявлення $h[Z_j]$ можуть бути отримані і для кожного з еталонів Z_j . Набір векторів $h[Z_j]$ утворює матрицю еталонних описів бази.

Значення вектору $h[Z_j]$ може виступити також критерієм якості класифікації на навчальній вибірці. Ідеальним класифікатором є той, для якого виконується умова: $\arg \max_j \{h_j[Z_k]\} = k$, тобто при класифікації еталона з номером k максимальне значення у векторі h досягається для компоненти з номером k . Чисельним критерієм за умови правильної класифікації є відношення числа дескрипторів, які за результатами оброблення віднесені до «правильного» класу, до загальної кількості s дескрипторів

$$\beta = \sum_{k=1}^J h_k[Z_k] / s. \quad (1.3)$$

Критерій (1.3) відображає частку правильно класифікованих дескрипторів із множини елементів навчальної вибірки. Ідеальним варіантом є значення $\beta = 1$.

Запропоноване класифікаційне правило \hat{g} фактично ототожнює клас об'єкта і клас дескриптора КТ, що виключає додаткові витрати на оброблення, викликані необхідністю переходу із кластерного простору до множини класів візуальних об'єктів, як запропоновано у [22].

1.2. Аналіз і вибір детекторів КТ

Зважаючи, що основу підходу, що розглядається у монографії, складають синтезовані описи зображень, виникає нагальна необхідність провести порівняльний аналіз можливостей, властивостей та характеристик детекторів для виявлення КТ та обчислення їх дескрипторів як базових даних для класифікації.

Останнім часом набули прикладного застосування такі детектори КТ, як ORB і BRISK [38, 39], які обчислюють дескриптор КТ у вигляді бінарного вектора з розмірністю, кратною ступеню двійки. Бінарне представлення значно прискорює процес зіставлення дескрипторів за рахунок можливості застосування двійкових операцій і відповідно спрощує апаратну реалізацію системи розпізнавання. Крім того, бінарна арифметика дає потенцію застосувати ефективний апарат оброблення двійкових даних і синтезувати нові підходи для визначення подібності у просторі дескрипторів при побудові правил класифікації.

Аналіз структури алгоритмів для реалізації детекторів ORB, BRISK показує, що вони менш вимогливі до обчислювальних ресурсів у порівнянні з іншими. Виграш у швидкості обчислень пояснюється спрощеною процедурою побудови дескрипторів, що дає за результатами тестування помітний виграш у швидкодії оброблення при порівняльній або кращій точності, ніж для традиційних детекторів SIFT і SURF [37]. Відзначається більш висока точність детектування з використанням BRISK на окремих тестових зображеннях у порівнянні із застосуванням SURF-дескрипторів.

Встановлення відповідності для дескрипторів ORB, BRISK здійснюється обчисленням відстані Хемінга, тобто підрахунком числа бітів, відмінних у двох бінарних векторах. Відповідні дії зводяться до побітової логічної операції XOR, що обчислюється ефективно. Модульність побудови ORB, BRISK дає можливість комбінувати побудову детектора КТ в

поєднанні з довільним способом визначення його дескриптора і навпаки, оптимізуючи бажану продуктивність вирішуваних завдань [3].

Таким чином, основні переваги методів ORB і BRISK полягають у забезпеченні більш високої продуктивності за рахунок спрощення процесу оброблення і побудови дескрипторів бінарного типу.

Історично раніше розроблено детектори SIFT (Scale Invariant Feature Transform) и SURF (Speeded Up Robust Features) [37]. Вони дають можливість сформуванню як множини координат КТ, так і визначити їх дескриптори. Метод SURF при цьому набув порівняльно до SIFT переважної популярності із-за поглибленого розроблення та забезпечує суттєво більшу швидкодію оброблення даних в процесі розпізнавання. Однак, тривають цілеспрямовані спроби створення більш простих способів детектування КТ, обчислення та порівняння дескрипторів, що забезпечують як достатній рівень інваріантності до спотворень, так і перевагу в швидкодії оброблення.

Обчислення, аналіз та оброблення КТ нами реалізовано у вигляді програмних засобів на основі модуля features2d до крос-платформної бібліотеки OpenCV 3.4, яку можна використати в операційних системах Windows, Linux, MacOS, Android, iOS. Програмна модель реалізована у вигляді засобу мовою C++ для використання в операційній системі Windows. Формат зображення – .png.

Бібліотека OpenCV написана мовою C++ і популярна за рахунок своєї відкритості та можливості безкоштовно використовуватися в навчальних та комерційних цілях. Також підтримується можливість використання бібліотеки у мовах програмування Python, Java, Ruby, Matlab, Lua та інших. Фактично, OpenCV – це набір типів даних, функцій і класів для обробки зображень алгоритмами комп'ютерного зору [3, 4, 86].

Зауважимо важливі для застосувань властивості детекторів ORB та BRISK. Так, метод ORB визначає значніші за розміром масиви OT у

порівнянні з SURF, однак, при цьому його дескриптори часто просторово скупчуються і не завжди відображають ключові особливості об'єкту [3].

У детекторі BRISK є можливість управління кількістю сформованих КТ. На число виявлених КТ впливають такі показники, як: поріг алгоритму FAST для різниці між інтенсивністю центрального пікселя та пікселями кола навколо нього; номер октави, що встановлює крок стиснення зображення; параметр масштабу для вибору околиці КТ; радіуси для BRIEF (у пікселях), де взято зразки навколо координат КТ; кількість точок для BRIEF при відборі зразків; мінімальний та максимальний пороги для пар точок при формуванні дескриптора та ін. [38, 39].

Дескриптори мають вид бінарного вектора розміром 256 для ORB та 512 для BRISK. Дескриптори кодуються у Open CV у вигляді типу uchar (unsigned char – тип даних в C++, використовують для зберігання символів, об'єм 8 біт, значення 0...255), а не в бітах. Вони зберігаються в матриці, де кількість рядків дорівнює числу виявлених дескрипторів, а число стовпців дорівнює 32 (256 бітів дескриптора трансформуються у 32 uchar).

Дескриптори ORB, BRISK (по два приклади) мають вигляд, зображений на рис. 1.2, де одиниця позначена чорним кольором, а нуль – білим [3, 4].

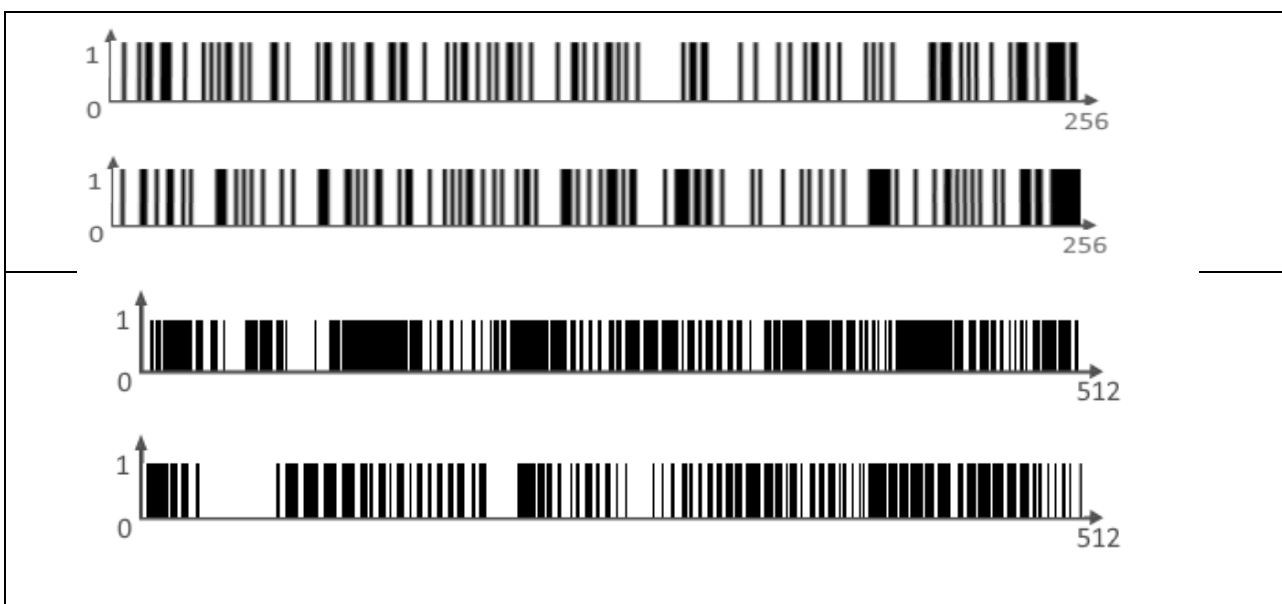


Рис. 1.2. Приклади значень дескрипторів ORB та BRISK

1.3. Грануляція значень множини дескрипторів

З метою забезпечення високої якості ідентифікації або розпізнавання довільних об'єктів важливо, наскільки описи еталонів розрізняються в рамках застосовуваного методу. Що значніша відмінність елементів описів або чим істотніше різняться їх склад, тим вище ймовірність правильного розпізнавання [24-30]. Наші дослідження показують, що показник швидкодії розпізнавання за рахунок трансформації системи ознак може бути поліпшений в десятки разів. Грануляція елементів множинного опису на підставі подібності його складових може забезпечити необхідне стиснення обсягу ознак, ефективну трансформацію простору, а також формування підмножин найбільш інформативних ознак з метою ефективного зниження обчислювальних витрат [34].

Метод SURF, застосовуваний для побудови описів, формує опис $Z \subset R^n$ як підмножину $R_1^n = \{z \mid z \in R^n, \|z\| \approx 1\}$ n -вимірних дійсних векторів, евклідова норма яких дорівнює одиниці [37]. На практиці ця умова реалізується в наближеному виді $R_1^n = \{z \mid z \in R^n, \|z\| \approx 1\}$.

Опис SURF може містити сотні векторів, представлених у формі з плаваючою комою, що загалом істотно сповільнює оброблення [28]. Завдання скорочення числа векторів передбачає побудову стисненого опису на основі відображення, варіантом якого є формування підмножини значно меншої потужності шляхом застосування процедури відбору (редукції) ознак із Z . У теорії розпізнавання це називають формуванням підмножини найбільш значущих або інформативних ознак [26, 27, 56].

Основним критерієм оцінювання подібності елементів є метрика ρ . Найчастіше в якості ρ використовують евклідову або манхеттенську метрики із R^n . Критерієм еквівалентності двох дескрипторів КТ є значення ρ , що не перевищує апріорної величини порога δ_z . Два дескриптори

вважаються еквівалентними, якщо виконано $\rho(z_1, z_2) \leq \delta_z$. Поріг δ_z визначають як відсоток (точність) від максимально можливого значення, він приймається рівним, наприклад, 1%, 3%, 5%, 10% [26]. В цілому вибір δ_z залежить також від застосовуваної процедури оброблення. Значення порогів можуть відрізнятися, наприклад, у процедурах навчання, при встановленні еквівалентності дескрипторів і при класифікації.

Розглянемо моделі формування, обробки і зіставлення компактних описів у вигляді підмножини унікальних дескрипторів, а також проаналізуємо результати проведеного комп'ютерного моделювання для дослідження ефективності модифікацій методів у порівнянні з відомими підходами [34].

Здатність гранулювання інформації – важлива властивість інтелектуальних систем розпізнавання, що використовують семантику для формалізації рішення задач. Відомі чіткий і нечіткий підходи гранулювання [40]. Грануляцію структурного представлення об'єктів розуміємо у двох аспектах: уявлення опису у вигляді вектора через перехід до мультимножини шляхом класифікації його елементів, а також як угруповання подібних елементів на основі властивостей унікальності або кластерного аналізу [34]. Принципова відмінність першого підходу полягає в тому, що при цьому апріорно задається базова множина, яка фактично визначає класи для елементів опису.

Інформаційною гранулою називають підмножину універсуму, на якому визначено відношення подібності (нерозрізненості, еквівалентності), гранула представляє собою об'єднання атомарних елементів [34, 40]. У результаті універсум або фіксований опис можна зобразити як множину гранул. Розроблено основи теорії міри і відношень на гранулах. Міра гранули найчастіше обчислюється як сума (інтеграл) значень функції приналежності елементів $d(A) = \sum_{a \in A} \mu_A(a)$, де $\mu_A(a)$ – значення функції приналежності елемента гранулі, $\mu_A(a) \in [0,1]$. Дії над гранулами здійснюються за законами

теорії множин. Гранули можуть включати одна одну, утворювати ієрархії. Виконання грануляції на практиці реалізується деяким методом і залежить від ряду параметрів. Можлива постановка питання про оптимальну грануляцію.

Для елемента z_k опису $Z = \{z_k\}_{k=1}^K$ визначимо число c_k повторювань:

$$c_k = \sum_{v=1, \overline{K}} 1(\rho(z_k, z_v) \leq \delta_z), \quad (1.4)$$

$$\text{де } 1(\rho(z_k, z_v) \leq \delta_z) = \begin{cases} 1, & \rho(z_k, z_v) \leq \delta_z, \\ 0, & \text{інакше.} \end{cases}$$

Таким чином, $c_k \in C_+$ – це число елементів Z , еквівалентних елементу $z_k \in Z$, C_+ – множина цілих невід'ємних чисел.

Розглянемо скінченну множину $Z = \{Z^i\}_{i=1}^J$ описів, що створюють базу еталонних зображень із J еталонів. Для кожного дескриптора $z_{ik} \in Z^i$ еталону з номером i обчислимо його індекси унікальності α_{ik}, β_{ik} :

$$\alpha_{ik} = c_{ik} / s_i ; \quad \beta_{ik} = c_{ik}^s / (s - s_i) ; \quad \alpha_{ik} \in [0,1], \quad \beta_{ik} \in [0,1], \quad (1.5)$$

де c_{ik} – число (1.4) подібних елементів для z_{ik} в i -му еталоні,

s_i – число дескрипторів i -го еталону,

c_{ik}^s – значення (1) для z_{ik} у базі еталонів без еталону з номером i ,

$s = \sum_i s_i$ – загальне число елементів описів бази Z .

Для елементів еталону обчислимо значення індексів (1.5), а також інтегральні показники $\sum_k \alpha_{ik}, \sum_k \beta_{ik}$. У загальному випадку $s_i \neq s_q$ при $i \neq q$. Індекс α_{ik} відображає ступінь повторювальності k -го елемента всередині i -го еталону, а індекс β_{ik} – рівень його повторювальності в описах решти об'єктів бази, крім i -го. Більш значущим для розпізнавання можна вважати

β_{ik} , чем ближче β_{ik} до нуля, тим більш унікальним є елемент серед решти елементів бази. Сума $\alpha_{ik} + \beta_{ik}$ відображає ступінь унікальності елемента в межах усієї бази. Інтегральні показники $\sum_k \alpha_{ik}, \sum_k \beta_{ik}$ характеризують спільну унікальність елементів еталону в межах бази. Зауважимо, що показники (1.4), (1.5) є результатом статистичного оброблення даних, яке є інструментом сучасної науки про дані.

Аналіз значень кортежу $\{z_{ik}, \alpha_{ik}, \beta_{ik}\}$ і інтегральних показників дає можливість практично оцінити ступінь розрізнення об'єктів бази при застосуванні структурної класифікації. Крім того, аналіз індексів унікальності дозволяє побудувати і застосувати процедури редукції ознак як всередині зразка, так і в межах бази. Фільтрація здійснюється шляхом побудови предиката «представництва» елемента у трансформованому описі. На основі обчислених характеристик можна виділити групи (гранули) рідко або часто повторюваних дескрипторів. Вони можуть бути основою побудови подібності об'єктів. Ефективність розпізнавання залежить від порогу δ_z і від числа сформованих унікальних ознак.

Запропонуємо конкретні схеми редукції ознак.

1. Еталони обробляються незалежно, визначається кортеж із t ознак з найменшими значеннями α_{ik} .

2. Для кожного зразка формується кортеж із t найбільш унікальних ознак серед інших еталонів, тобто з найменшими значеннями β_{ik} .

Виграш у швидкодії розпізнавання для цих схем у порівнянні із традиційним зіставленням множин векторів становить s_i / t . Трансформація простору ознак при цьому здійснюється на попередньому етапі і не впливає на час розпізнавання. Друга схема більш перспективна в плані впливу на ефективність, тому що критерій β_{ik} відображає ступінь відмінності еталонів. Ознаки з високим значенням α_{ik} (часто повторювані всередині опису) також

мають інформаційне навантаження і позначаються на результатах розпізнавання.

Визначення поняття і значень статистичних індексів унікальності призводить до думки про можливість подання елементів бази описів у деякій системі класів, що гранулює елементи в групі i , як результат, сприяє значному скороченню часу оброблення [34]. Класи можуть бути сформовані, наприклад, на основі характеристик α_{ik}, β_{ik} . Одним із варіантів є «проріджування» множини ознак.

У результаті класифікації по базі $U = \{u_1, u_2, \dots, u_i, \dots\}$ різних за значенням елементів отримаємо уявлення еталона як мультимножини. Належність точки z мультимножині A встановлюється шляхом обчислення характеристичної

функції $\chi_A(z) = \begin{cases} 1, & z \in A, \\ 0, & z \notin A \end{cases}$, де перевірка $z \in A$ виконується порівнянням z з

елементами U на основі $\rho(z, u_i)$ або іншої міри. Клас u_* елемента z визначимо шляхом дискретної оптимізації

$$u_* = \arg \min_{u_i \in U} \rho(z, u_i), \quad \rho(z, u_*) \leq \delta_z, \quad (1.6)$$

на множині U . Класифікація (1.6) включає перевірку умови $\rho(z, u_*) \leq \delta_z$, в результаті окремі елементи взагалі можуть бути відкинуті. На основі (3) при m базових класах для еталона $Z(j)$ отримаємо векторне подання

$$Y(j) = \{\gamma_1, \gamma_2, \dots, \gamma_i, \dots, \gamma_m\}, \quad \gamma_i \in C_+, \quad (1.7)$$

де γ_i – кратність входження i -го базового елемента.

Класифікацією (1.6) і формуванням опису (1.7) здійснена трансформація $Z \rightarrow Y$ із множинного подання до векторного опису Y скінченної розмірності m з компонентами із C_+ . Результат класифікації і

трансформації опису визначається параметром δ_z . При дії завад реалізація (1.6) додатково забезпечує фільтрацію перешкод.

Побудовою (1.7) здійснена грануляція структурного опису у вигляді кластерів близьких елементів. Це інтерпретується як класифікація, вона відрізняється від традиційної кластеризації, тому що «представники» кластерів задаються апріорно базовою множиною U . Схема класифікації на основі U схожа з розкладанням по системі ортогональних функцій. Основна відмінність полягає в тому, що для елементів ортогональність не потрібна, тому що у відновленні елементів початкової системи ознак немає необхідності.

Розпізнавання об'єкту при поданні мультимножиною здійснюється шляхом обчислення та мінімізації відстані, наприклад,

$$\rho(A, B) = \frac{\sum_i w_i |\gamma_A(u_i) - \gamma_B(u_i)|}{\sum_i w_i \max[\gamma_A(u_i), \gamma_B(u_i)]}, \quad (1.8)$$

де w_i – вагові коефіцієнти (часто приймають $\sum_i w_i = 1$),

$\gamma_A(u_i), \gamma_B(u_i)$ – кратності при поданні множин A, B у проекції базового елемента $u_i \in U$.

Таким чином, у результаті проведення грануляції шляхом уявлень (1.6) або (1.7) для елементів Z ми отримали дворівневу систему класифікації, де клас об'єкта визначається як

$$j = K_2[K_1[Z, U]], \quad (1.9)$$

де K_1 – класифікація на рівні елементів опису, K_2 – класифікація трансформованого опису у базі Z .

Ключовим моментом, що впливає на результат класифікації із застосуванням запропонованого методу, є створення базової множини

$U = \{u_i\}_{i=1}^m$. Множина U – це якийсь тезаурус, ознаки опису «знаходять своє місце» в уявленні $\Upsilon(j)$. Елементи, віднесені до u_i , вважаємо об'єктами-синонімами, вони при розпізнаванні формально не відрізняються.

Нехай s – вихідне число ознак, обчислення індексів унікальності і класів ґрунтується на аналізі симетричної матриці відстаней $H = \|h_{i,k}\|_{s \times s}$, де $h_{i,k} = \rho(z_i, z_k)$. Відстані відображають зв'язок між елементами у вигляді ступеня еквівалентності. Для кожного z_i можна визначити характеристики сумарної і середньої зв'язності

$$d_i(z_i) = \sum_{k=1}^s h_{i,k}, \quad \bar{d}_i = \frac{1}{s-1} \sum_{k=1, k \neq i}^s h_{i,k}, \quad (1.10)$$

де використовуються елементи i -го рядка (стовпчика) матриці H . Чим менше значення (1.10), тим сильніше i -й елемент пов'язаний з іншими із опису.

Далі здійснимо упорядкування елементів у порядку зниження рівня зв'язності, наприклад, у вигляді $d_{\min} = d_1 \leq d_2 \leq \dots \leq d_s = d_{\max}$. У результаті створюється порядок, відповідно до якого вибираються найбільш інформативні (в плані найбільшого) елементи. Інший варіант аналізу – підрахунок кількості елементів, з якими пов'язаний аналізований елемент відповідно до (1.5). Тут будується свій порядок елементів. В обох схемах перевіряється умова $r_{i,k} \leq \delta_r$, тобто аналізуються тільки вагомі зв'язки.

Моделі оброблення (1.5), (1.10) служать основою побудови множини U і встановлення числа m , від якого обернено пропорційно залежить виграш у швидкодії, пов'язаний із впровадженням стисненням простору ознак. Іншим способом формування є редукція ознак за значеннями характеристик (1.5), (1.10) або застосування кластеризації на множині елементів бази. Елементи U тут вибираються як представники або центри кластерів. Схема

перетворення опису при формуванні інформативних ознак або векторного уявлення приведена на рис. 1.3.

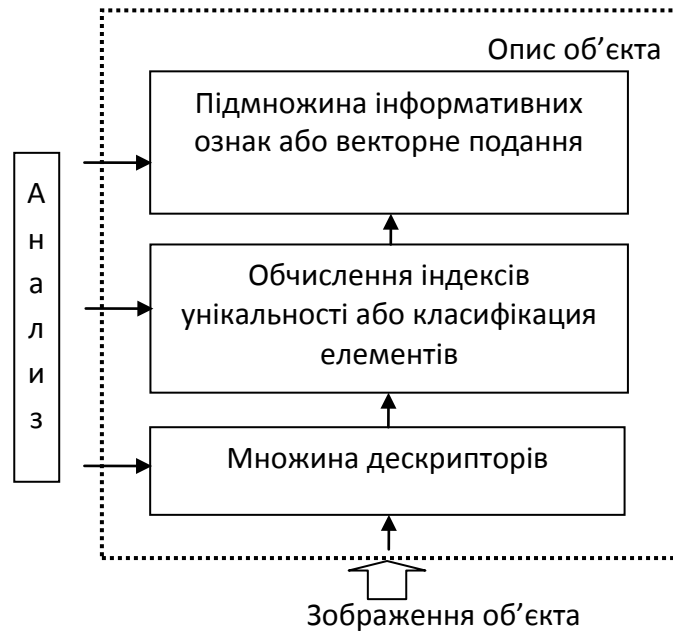


Рис. 1.3. Етапи оброблення опису

Наші експерименти показали, що для різних баз зображень при порозі $\delta_z < 0,48$ для значної частини дескрипторів SURF (більше 80%) не визначається жодної схожої ознаки, а при $\delta_z > 1,6$ всі ознаки стають еквівалентними між собою. В цих обох граничних ситуаціях неможливе формування підмножини унікальних ознак. Для отримання множини унікальних даних найбільш прийнятні значення порогів 0,64 і 0,8. На рис. 2 наведені зображення 300×300 пікселів із досліджуваної бази облич. Число дескрипторів SURF в їх описах склало 96, 89, 98.

Табл. 1.1 містить значення індексів унікальності (1.5) для 10-ти перших точок опису та індекс для опису зображення рис. 2, а.

Розглянуті 10 КТ мають досить близькі значення (1.5). В цілому діапазон значень для опису рис. 1.4, а становить від 0,1 до 0,7. Рис. 1.5 містить гістограму повторень КТ в описі зображення рис. 1.4, а. Як бачимо із рис. 1.5, об'єктивно існує група унікальних ознак, яку можна взяти за основу

при формуванні компактного опису. Табл. 1.2 містить номери і значення індексів для 10-ти найбільш унікальних КТ опису. На рис. 1.6 показана множина сформованих КТ і підмножина із 10 унікальних КТ.

Критерій унікальності є ключовим для побудови стисненого опису, а число КТ може бути скорочено відразу в 10 разів. Моделювання показало, що ймовірність правильного розпізнавання на основі стисненого опису у вигляді 10-ти унікальних КТ при відсутності перешкод не знижується. Зате час розпізнавання в нашому модельному експерименті знизився в 2,6 рази. Скорочення обчислювальних витрат безпосередньо залежить від розміру бази. Наприклад, при числі зображень 30 виграш перевищує 100 раз. Для інших варіантів оброблення, наприклад, якщо виділити унікальні ознаки на розпізнаваному зображенні, то виграш для трьох еталонів уже становить приблизно 20 разів.

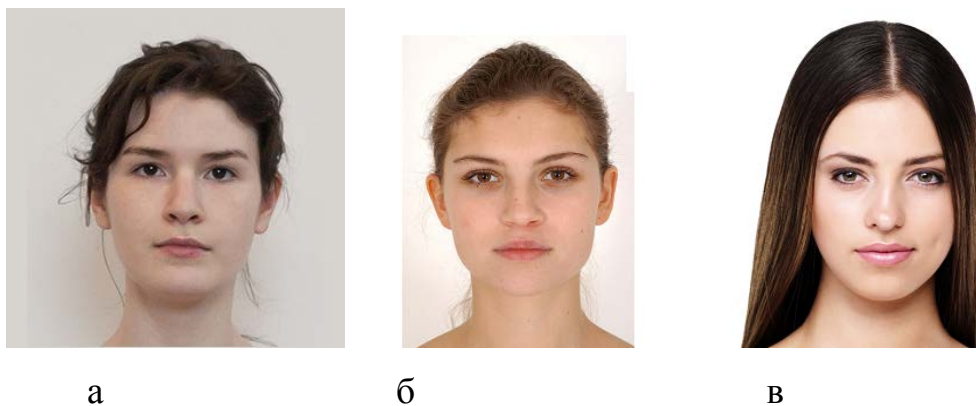


Рис. 1.4. Приклади зображень жіночих облич

Таблиця 1.1. Значення індексів унікальності

№	1	2	3	4	5	6	7	8	9	10	Сума
α_{ik}	0,34	0,52	0,4	0,34	0,26	0,48	0,55	0,4	0,48	0,67	39,46
β_{ik}	0,24	0,45	0,29	0,4	0,32	0,49	0,48	0,5	0,48	0,65	36,36

Проведені експерименти показали, що число помилкових збігів для системи унікальних КТ зображень істотно менше, ніж для повних описів.

Наприклад, при порозі $\delta_z = 0,64$ кількість помилкових збігів знижується з 96-97% до 10-40%, а при порозі $\delta_z = 0,48$ – з 57% до 10%.

Проведена також класифікація в умовах спотворення зображень адитивним шумом з гаусовим розподілом. При впливі шуму число КТ, як правило, збільшується (в нашому експерименті – в 3,3 рази). При проведенні розпізнавання на основі множин унікальних КТ еталонів кількість помилкових відповідностей скоротилося з 58% практично до 0. При цьому ймовірність розпізнавання не знижується, тому що число відповідностей з шуканим об'єктом становить 80%. Близькі результати отримані також і для бази з 54 зображень гербів країн Європи (рис. 1.7).

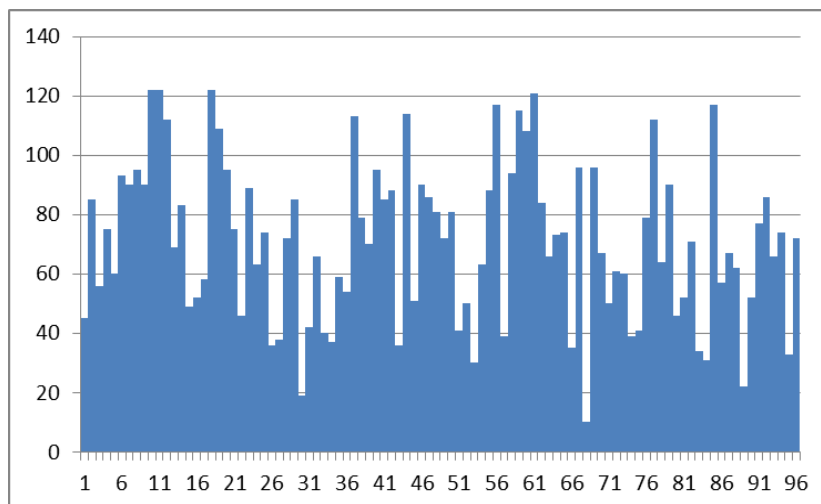


Рис. 1.5 Гістограма повторювання значень КТ в описі зображення рис. 2, а з зображеннями 2, б і 2, в.

Таблиця 1.2. Значення індексів унікальності для відібраних КТ

№	67	29	88	83	82	21	50	70	94	14
α_{ik}	0,1	0,11	0,11	0,14	0,17	0,22	0,23	0,23	0,23	0,24
№	67	29	88	52	83	94	82	65	25	42
β_{ik}	0,05	0,1	0,11	0,16	0,16	0,17	0,18	0,18	0,19	0,19

Гістограми рис. 1.8 показують число подібних КТ зображення рис.1.4, а з 10 унікальними ознаками всередині нього (рис. 1.8, а) і в рамках всієї бази (рис. 1.8, б) без застосування порога δ_z . Вид цих гістограм підтверджує можливість розпізнавання на основі класифікації ХП, де базовою множиною виступає 10 унікальних КТ.

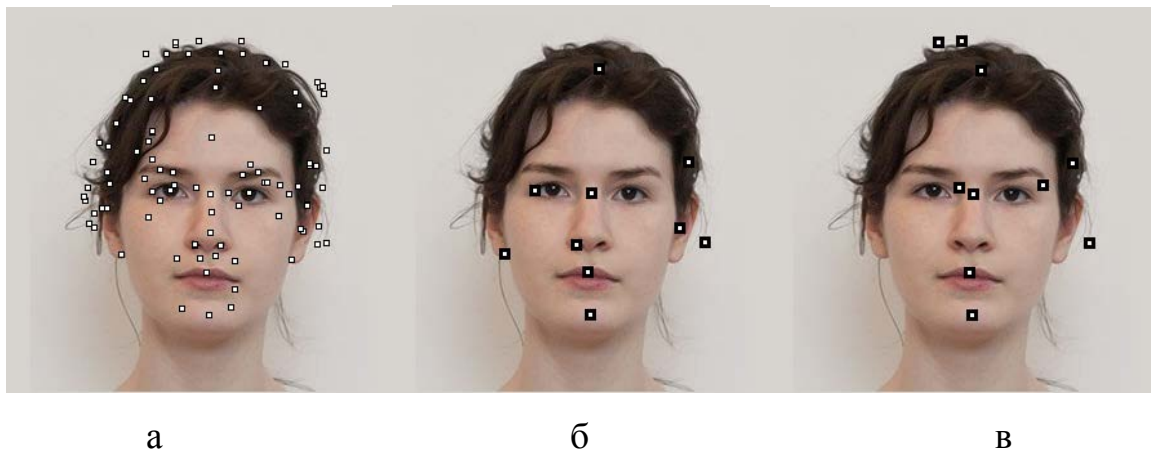


Рис. 1.6. Зображення 2, а з множиною КТ (а), 10 найбільш унікальних ознак (б), 10 найбільш унікальних ознак в базі із трьох зображень (в).



Рис. 1.7. Приклади з бази зображень гербів країн Європи

Теоретично обговорювані методи можуть здійснювати розпізнавання навіть за однією унікальною ознакою, якщо вхідне зображення не піддавалося спотворенням чи дії перешкод. У той же час для прикладних задач число унікальних ознак повинно бути, як правило, більше 10 в залежності від вимог застосування. Практичне число класів для дескрипторів КТ, як показало моделювання, знаходиться в межах 10-20.

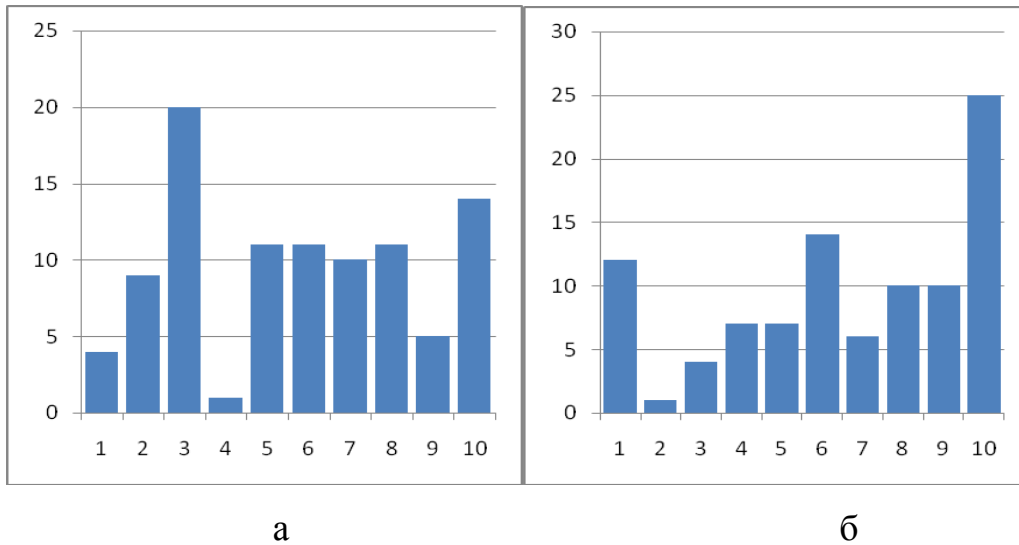


Рис. 1.8. Гістограми подібності КТ зображення 2, а

Висновки

1. Побудова методів структурної класифікації зображень ґрунтується на описі як значеннях множини дескрипторів ключових точок. Властивості цієї множини даних у аспекті розрізнення визначають результативність класифікації у межах бази еталонів.

2. Вибір типу детектору ключових точок визначається вимогами до апаратних та програмних засобів реалізації класифікатора. Застосування бінарних дескрипторів спрощує обчислення і сприяє зниженню затрат.

3. Аналіз статистичних показників унікальності для елементів описів дає можливість забезпечити необхідний рівень розрізненості даних у процесі класифікації, а також суттєво скорочує обсяги обчислювальних витрат.

4. Проведені експерименти з моделювання методу класифікації на підставі грануляції даних показали підвищення швидкодії оброблення в десятки разів, що підтверджує практичну цінність отриманих результатів дослідження.

2. Статистичний аналіз даних для описів у вигляді системи кластерів

2.1. Класифікація на засадах байєвської теорії прийняття рішень

Формальною основою структурного розпізнавання є база описів зображень еталонів (алфавіт класів), що подана у вигляді сукупності дескрипторів КТ як множини $Z = \{Z^j\}_{j=1}^J$ [28, 65].

Виконаємо на множині Z деяке розбиття $Z = M = \{M_i\}_{i=1}^k$, $M_i \cap M_d = \emptyset, M_i \neq \emptyset$. Вважаємо $z \in M_i$ еквівалентними між собою. Маємо дві системи класів $\forall z \in Z: \{Z^j\}$ – для зображень еталонів та $\{M_i\}$ – для кластерів. Розбиттям $\{M_i\}_{i=1}^k$ опис Z^j еталону трансформується у вектор цілих чисел

$$h[Z^j] = (h_1^j, h_2^j, \dots, h_i^j, \dots, h_k^j). \quad (2.1)$$

Подання (2.1) – це образ еталону в кластерному вигляді, а множина векторів (2.1) бази із J зразків має вигляд матриці $H[Z] = \{\{h_i^j\}_{j=1}^J\}_{i=1}^k$. Рядок матриці H відображає кластерний опис для окремого еталону, а стовпець – зміст кластера із числа рівноцінних елементів різних еталонів.

Сучасні структурні методи зводяться до оцінювання подібності дескрипторних описів аналізованого і еталонного зображень, що представлені у вигляді множин КТ [28, 66]. Значення отриманої оцінки відображає ступінь релевантності двох інформаційних одиниць, які є описами об'єктів у побудованому просторі кількісних ознак.

Множина дескрипторів КТ у вигляді сукупності числових векторів формується за допомогою детекторів [23, 37]. Одним із шляхів продуктивного співставлення базової множини КТ з аналізованими об'єктами є побудова описів еталонів на підставі кластеризації, що значно знижує розмірність досліджуваних даних за рахунок формування їх

групованої просторової структури [23, 67]. Результатом кластеризації є образ об'єкта (2.1), який можна розглядати як певну комбінаторну конфігурацію, організовану із елементів базової множини центрів кластерів $m = \{m_1, m_2, \dots, m_k\}$, $m_i \in M_i$, на підставі сформованої кластерної структури.

Практичне впровадження методів структурного розпізнавання потребує застосування механізму оцінювання їх результативності для випробуваних прикладних зразків даних. Одним із таких засобів є апарат баєсовської теорії прийняття рішень, що є фундаментальним статистичним інструментарієм розпізнавання образів [43, 68]. Баєсовський класифікатор вирішує задачу дискримінантного аналізу і функціонує в термінах апостеріорної ймовірності віднесення об'єкта до класу [45, 69]. Ймовірнісний підхід також дає можливість здійснити попереднє оцінювання результативності структурного розпізнавання в межах прикладної бази зображень.

Під класифікацією розуміємо відображення множини описів об'єктів $\{O\}$ в скінченну множину номерів еталонів $\{1, \dots, J\}$, що здійснюється шляхом віднесення опису $O = \{o_i\}$ невідомого візуального об'єкта до однієї з еталонних множин $Z = \{Z^j\}_{j=1}^J$ (рис. 2.1). Класифікацію базуємо на теоретико-множинному описі даних, що передбачає кластерну модель на підставі статистичного подання (2.1).



Рисунок 2.1. Схема структурного розпізнавання

Процес класифікації будемо здійснювати порівняно за двома підходами [13-16]:

- 1) шляхом зіставлення інтегральних характеристик образу виду (2.1) [23];
- 2) застосуванням баєсовського класифікатора [45].

Розглянемо сутність обох методів класифікації більш детально.

Метод 1.

1. Для побудови кластерного опису для $Z = \{Z^j\}_{j=1}^J$ нормуємо матрицю $H[Z] = \{\{h_i^j\}_{j=1}^J\}_{i=1}^k$ в рядках, отримуємо:

$$\alpha^i[Z^j] = \{\{h_i^{j*}\}_{j=1}^J\}_{i=1}^k = \{\{h_i^j / s_j\}_{j=1}^J\}_{i=1}^k. \quad (2.2)$$

2. Віднесемо кожний елемент $o_l \in O$ об'єкта до кластеру $M_i \subseteq M$ у відповідності з конкурентним правилом:

$$o_l \rightarrow M_i \mid \arg \min \rho(o_l, m_d) = i. \quad (2.3)$$

Для фільтрації помилкових елементів, які можуть бути отримані в результаті реалізації (2.3), виконаємо верифікацію мінімуму $m_i: \rho(o_l, m_i) \leq \delta$. Якщо нерівність не виконується, елемент o_l не відносимо до жодного з кластерів.

3. За результатом виконання кроку 2 формуємо опис (2.1) об'єкта $O = \{o_l\}: O = (h_1, h_2, \dots, h_k)^o$.

4. Обчислюємо ступінь r_j релевантності нормованих кластерних описів як відстань $r_j = \beta(\alpha^i[O], \alpha^i[Z^j])$ для об'єкта та рядків матриці еталонів у просторі R^k .

5. Віднесемо об'єкт O до класу $d \in \{1, \dots, J\}$ за правилом: $d = \arg \min_j r_j$.

Перевіримо цінність мінімуму релевантності $r_d : r_d \leq \varepsilon$, де ε – встановлений поріг. У разі порушення умови клас об'єкта вважаємо невизначеним через відсутність значущої відповідності у наявному еталонному просторі.

Метод 2.

1. Виходячи з початкових кластерних описів еталонів $Z = \{Z^j\}_{j=1}^J$ і об'єкта $O = \{o_l\}$, обчислимо за формулою Баєса апостеріорні ймовірності $P(Z_j / O)$ віднесення об'єкта послідовно до кожного з описів $h[Z^j]$, $j \in \{1, \dots, J\}$ [43, 68]:

$$P(Z_j / O) = \frac{P(O / Z^j) \cdot P(Z^j)}{\sum_{j=1}^J P(O / Z^j) \cdot P(Z^j)}, \quad (2.4)$$

де $P(O / Z^j)$ – апіорна ймовірність належності об'єкта до еталону Z^j , $P(Z^j)$ – ймовірність появи Z^j , яку для спрощення вважаємо рівноймовірною.

Дискретний характер проблеми віднесення об'єкта $O = (h_1, h_2, \dots, h_k)^o$ до еталону Z^j в кластерному поданні призводить до обчислення набору ймовірностей узагальненого гіпергеометричного розподілу [70]:

$$P(O / Z^j) = \prod_{i=1}^k C_{h_i^j}^{h_i} / C_{s_j}^{s_j}. \quad (2.5)$$

Зауважимо, що формула (2.5) має сенс лише для невід'ємних цілих значень параметрів, які задовольняють умовам:

$$h_i \leq h_i^j, \quad s_j \leq s, \quad i = 1, \dots, k, \quad j = 1, \dots, J. \quad (2.6)$$

Обмеження (2.6) у практичних розрахунках можуть потребувати певного корегування вхідних даних, яке можна здійснювати, в першу чергу, шляхом пропорційного змінювання характеристик досліджуваного об'єкта. Коригування кластерного подання об'єкта при невиконанні обмежень (2.6) пропонуємо здійснювати з округленням результату до цілого числа за формулою:

$$h'_i = \lambda_{\text{кор}} \cdot h_i, \quad i = 1, \dots, k.$$

Розрахунок здійснюється для $i \in \{1, \dots, k\}$, для яких порушується умова (2.6).

Вказані корегування загалом можуть вплинути на числові результати розрахунків, але принципово не змінюють рішення щодо факту найімовірнішого віднесення об'єкта до одного з еталонів. Відмітимо також, що при застосуванні формули (2.5) можливе виникнення складнощів при невеликих значеннях h_i^j . Це може означати надмірно дрібне розбиття на кластери, якого можна запобігти шляхом їх укрупнення.

2. Віднесемо об'єкт O до класу $d \in \{1, \dots, J\}$ за правилом:

$$d = \arg \max_j P(Z_j / O).$$

Як бачимо, перший метод обчислює ступінь релевантності шляхом дефініції подібності кластерних описів зображень об'єкта і еталонів, а другий – на підставі визначення значень ймовірності віднесення опису до відповідного класу. Оцінювання результативності процедур структурного розпізнавання в межах конкретної прикладної бази зображень здійснимо на підставі порівняння відповідних числових результатів – матриці відстаней між еталонами у методі 1 та ймовірностей віднесення еталону до класу у методі 2.

Розглянемо приклад реалізації методу 2 за допомогою експериментальної бази з числом еталонів 4 в поданні із 10 кластерів, до якої

у [23] було застосовано метод 1. Проаналізуємо особливості використання методу 2, а також проведемо порівняльний аналіз результатів, одержаних за обома методами.

Використаємо кластерне подання у вигляді таблиці 2.1 [23].

Таблиця 2.1. Кількість КТ еталонів у кластерному представленні

Еталон	Кластер									
	M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8	M_9	M_{10}
Z^1	7	9	9	9	14	15	6	11	11	9
Z^2	8	11	8	16	3	14	13	9	11	7
Z^3	15	7	3	13	9	18	7	8	11	9
Z^4	10	8	8	13	13	12	7	11	11	7

Процес знаходження апостеріорних ймовірностей $P(Z_j / O)$ віднесення об'єкта до описів $h[Z^j]$, $j \in \{1, \dots, J\}$ включає застосування формули (2.5) та передбачає виконання умов (2.6). Продемонструємо реалізацію підходу, взявши в якості об'єктів самі еталони.

Так, кластерне представлення об'єкта O^1 , яким вибрано еталон Z^1 , має перевищення кількості елементів у кластерах $M_2, M_3, M_5, M_6, M_8, M_{10}$ у порівнянні з кластерним поданням еталонів Z^2, Z^3, Z^4 . Аналізуючи співвідношення між кількостями елементів відповідних кластерів, отримуємо найменший з коефіцієнтів пропорційності, рівний 0,21, що приводить до відкорегованого кластерного опису об'єкта O^1 , що задовольняє (2.6):

$$O_{кор}^1 = (2, 2, 2, 2, 3, 3, 1, 2, 2, 2).$$

Таке коригування забезпечує можливість застосування традиційного підходу щодо розрахунку апостеріорних ймовірностей (2.5). При цьому вираз (2.5) обчислюється безпосередньо і, наприклад, для еталона Z^1 приймає вигляд:

$$P(O_{кор}^1 / Z^j) = C_7^2 C_9^2 C_9^2 C_9^2 C_{14}^3 C_{15}^3 C_6^1 C_{11}^2 C_{11}^2 C_9^2 / C_{100}^{21}.$$

Обираючи об'єктами еталони Z^2, Z^3, Z^4 , аналогічно отримуємо відкориговані кластерні описи:

$$O_{кор}^2 = (3, 4, 3, 6, 1, 5, 5, 3, 4, 3), \quad O_{кор}^3 = (5, 2, 1, 4, 3, 6, 2, 3, 4, 3), \quad O_{кор}^4 = (2, 2, 2, 3, 3, 3, 2, 3, 3, 2).$$

Вважаючи появу еталонів рівноймовірною, що у нашому випадку означає $P(Z^j) = 0,25$, розраховуємо за формулою (2.4) апостеріорні ймовірності віднесення $O_{кор}^1, O_{кор}^2, O_{кор}^3, O_{кор}^4$ послідовно до кожного з описів $h[Z^j], j \in \{1, \dots, 4\}$. Результати обчислень занесемо у таблицю 2.2.

Таблиця 2.2. Апостеріорні ймовірності віднесення коригованих описів об'єктів

Еталон	Об'єкт			
	$O_{кор}^1$	$O_{кор}^2$	$O_{кор}^3$	$O_{кор}^4$
Z^1	0,536	0,001	0,014	0,389
Z^2	0,005	0,993	0,001	0,014
Z^3	0,069	0,001	0,927	0,063
Z^4	0,389	0,005	0,058	0,533

У відповідності до методу 2 (крок 2) згідно з баєсовською теорією відносимо аналізований об'єкт до еталону, який має найбільше значення апостеріорної ймовірності. Результат класифікації підтверджує коректність роботи розробленого методу 2, оскільки кожний із тестових об'єктів розпізнаний правильно.

Зауважимо, що отримані значення ймовірностей мають суттєві відмінності: перший і четвертий еталони розпізнано не так впевнено (із значно нижчими показниками ймовірностей), ніж другий і третій. Ці ж висновки, що підкреслюють особливості розглянутої експериментальної бази даних, отримано нами у роботі [23], що підкреслює близькість обох методів.

Зауважимо, що схожість теоретичного підґрунтя статистичних методів та методів порівняння з еталоном для бінарних зображень доведено у [68].

Адекватність проведених розрахунків можна також підтвердити значенням сукупного критерію близькості кожного з векторних описів у аналізованій базі еталонів [28]:

$$\gamma_d = \sum_{\substack{j=1 \\ j \neq d}}^J \sum_{i=1}^k |h_i^j - h_i^d|. \quad (2.7)$$

Чим менше значення γ_d , тим ближче цей еталон до сукупності решти еталонів. Відповідно, і класифікується він гірше. У нашому випадку за таблицею 1 критерій (2.7) має значення $\gamma_1 = 82$, $\gamma_2 = 104$, $\gamma_3 = 96$, $\gamma_4 = 70$. Це свідчить про значнішу наближеність першого і четвертого еталонів до сукупності еталонів загалом, ніж другого і третього, що можна вважати додатковим поясненням відмінностей значень ймовірностей у таблиці 2.2. Як бачимо, еталон з номером 4 різниться в базі з найменшою ефективністю. Критерій (7) аналогічно матриці відстаней [68] має універсальне призначення щодо оцінювання ефективності довільних систем кількісних ознак.

Зауважимо, що кластерне подання еталонів і досліджуваного об'єкту можна розглядати як мультимножини, зважаючи, що центри кластерів виступають базовими елементами мультимножини. Отже, критерій (2.7) з цієї точки зору має сенс відстані між фіксованою мультимножиною та сукупністю мультимножин бази еталонів, що може безпосередньо застосовуватися в прикладних задачах розпізнавання візуальних об'єктів [66].

Тепер зупинимося на результатах розрахунків щодо реалізації запропонованого методу 2 на прикладі експериментальної бази зображень гербів міст України (Дніпро, Київ, Харків, Кременчук), сформованої на

підставі застосування детектору SURF [65]. Кластерне подання має вигляд табл. 2.3.

Таблиця 2.3. Кількість СО еталонів у кластерному представленні

Еталон	Кластер							
	M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8
Z^1	40	75	24	9	5	61	5	34
Z^2	10	91	38	27	5	66	4	63
Z^3	102	113	20	17	4	14	3	69
Z^4	83	51	57	10	5	41	5	61

Результати обчислень, що представлені у табл. 2.4, свідчать про достатньо високий рівень ефективності розпізнавання, що можна пояснити апріорними суттєвими відмінностями між еталонами аналізованої бази у використовуваному просторі ознак.

Таблиця 2.4. Апостеріорні ймовірності розпізнавання коригованих об'єктів

Еталон	Об'єкт			
	$O_{кор}^1$	$O_{кор}^2$	$O_{кор}^3$	$O_{кор}^4$
Z^1	0,9999	0,00002	0,0001	0,0026
Z^2	0,0000	0,99998	0,0000	0,0000
Z^3	0,0000	0,00000	0,9992	0,0001
Z^4	0,0001	0,00000	0,0007	0,9973

Суттєва відмінність між еталонами в табл. 2.4 підтверджується також розрахунками за критерієм (2.7), який у ситуації різної кількості елементів в еталонах пропонуємо застосовувати через перехід до нормованого кластерного представлення [65]. Дійсно, критерій (2.7) приймає значення $\gamma_1 = 134$, $\gamma_2 = 150$, $\gamma_3 = 154$, $\gamma_4 = 155$, що не дає підстав стверджувати про наявність явно вираженої близькості одного з еталонів до сукупності решти еталонів бази.

Проаналізуємо тепер результати застосування розробленого методу 2 до експериментальної бази морських ссавців [23], структурний опис якої отримано за допомогою детектора SURF, вхідне кластерне подання бази містить таблиця 2.5. Результати обчислень наведемо в таблиці 2.6.

Таблиця 2.5. Кількість СО еталонів для бази морських ссавців

Еталон	Кластер				
	M_1	M_2	M_3	M_4	M_5
Z^1	10	11	3	7	8
Z^2	19	14	0	20	16
Z^3	7	12	4	26	5
Z^4	8	4	3	6	7

Таблиця 2.6. Апостеріорні ймовірності розпізнавання коригованих об'єктів

Еталон	Об'єкт			
	$O_{кор}^1$	$O_{кор}^2$	$O_{кор}^3$	$O_{кор}^4$
Z^1	0,565	0,026	0,007	0,080
Z^2	0,253	0,952	0,090	0,130
Z^3	0,008	0,000	0,900	0,000
Z^4	0,174	0,021	0,003	0,789

Як бачимо із розрахункової табл. 2.6, результат класифікації реальних зображень також підтверджує коректність роботи розробленого методу 2. Зауважимо, що відповідні обчислення для 1-го, 3-го і 4-го еталонів тут здійснювалися без урахування значень 3-го кластера, що пов'язано із необхідністю дотримання обмежень (2.6) для коректного застосування формули (2.5). Тут можна вважати, що цей кластер містить настільки незначну у порівнянні з іншими кількість елементів, що не може принципово вплинути на остаточний результат щодо правильного розпізнавання об'єкту.

Зауважимо, що обчислення зіставних значень критерію (2.7) в умовах різного числа СО в описах еталонів та можливих нульових значень у поданні окремих еталонів потребує додаткового коригування вхідних параметрів h_i^j для виразу (2.7). Необхідне попереднє коригування може бути виконане шляхом вирівнювання кількості СО еталонів фільтрацією чи випадковим вибором, виключенням із аналізу окремих несуттєвих за значенням кластерів чи нормуванням через приведення цих параметрів до фіксованого діапазону.

Із проведених розрахунків для змодельованих та реальних описів візуальних об'єктів можна зробити узагальнюючий умовивід про універсальні властивості запропонованого методу 2 щодо прийняття рішення про клас об'єктів на підставі їх структурного опису, трансформованого до кластерного вигляду. Ключовим критерієм, який можна застосувати для верифікації розрахунків, є значення (2.7).

Провідним результатом дослідження є підтвердження та перевірка розрахунком фундаментального зв'язку методів порівняння з еталоном та статистичного підходу у розпізнаванні образів стосовно структурних описів у вигляді множини характерних ознак зображень, які результативно представлено у кластерним апаратом. Ці підходи при розпізнаванні можуть бути використані одночасно або для контролю результату прийняття рішень. З іншого боку, запропонований у роботі більш простіший в аспекті обчислювальних витрат статистичний підхід на підставі баєсовських оцінок може застосовуватися для попередніх розрахунків ефективності розпізнавання без проведення затратних експериментів з програмного моделювання.

Засвідчено ефективність розробленого методу обчислення ймовірнісних оцінок якості розпізнавання для прикладних баз зображень. Результат класифікації продемонстрував коректність застосування методу, оскільки кожний із тестових об'єктів у декількох розглянутих базах зображень розпізнаний правильно.

Здійснено порівняння отриманих кількісних результатів обчислень із експериментальними даними комп'ютерного моделювання.

Порівнюючи результати розпізнавання за двома аналізованими підходами – шляхом визначення інтегральних характеристик для структурного опису образу та застосуванням баєсовського класифікатора на підставі кластерного подання – отримано порівнянні результати класифікації. Числовими результатами є матриці відстаней між еталонами у першому методі та ймовірності віднесення еталону до класу – у другому.

Дискретна природа постановки проблеми і, відповідно, обчислення набору ймовірностей узагальненого геометричного розподілу в процесі застосування формули Баєса у розробленому методі 2 призводить до накладення певних обмежень щодо значень параметрів образу, задоволення яких може потребувати додаткового коригування вхідних даних – як на етапі формування кластерної структури еталонів і аналізованого об'єкту, так і на етапі роботи із сформованими структурними описами.

2.2. Класифікація на підставі статистичних мір релевантності

Задача класифікації у векторному просторі може бути вирішена потужними засобами математичної статистики, призначення якої як раз і полягає у добуванні якомога більш повної інформації про аналізовані явища, виходячи зі обмеженого обсягу даних спостережень. Зауважимо [2, 8, 9], що існує два принципово різних підходи статистичного оцінювання – баєсовський і частотний (зокрема, метричний). Перший інтерпретує випадковість як міру незнання, що закладено в ідеї переходу від апіорних знань до апостеріорних з урахуванням здійснених спостережень. Другий розглядає випадковість як об'єктивну невизначеність і спирається виключно на наявні спостереження, як правило, на частоти появи досліджуваних подій,

отже, не пов'язаний ні з якими апіорними припущеннями щодо розподілу досліджуваних величин.

Кожний із зазначених підходів може бути підвалиною для побудови мір релевантності для векторних описів об'єктів [6, 7]. Ці міри для класифікації зображень мають певні особливості, зокрема, як переваги, так і окремі недоліки, що потребує осмислення і докладного аналізу із тестуванням.

2.2.1. Баєсовський класифікатор

Застосуємо апарат баєсовської теорії прийняття рішень [68, 71]. Обчислення апостеріорних ймовірностей віднесення опису об'єкта до множини еталонів є підставою для безпосереднього здійснення класифікації. Основна ідея апарату Баєса полягає у ідентифікації аналізованого об'єкту з еталоном, що має найбільше значення апостеріорної ймовірності.

Виходячи з кластерних описів еталонів $Z = \{Z^j\}_{j=1}^J$ і об'єкта $O = \{o_i\}$, за формулою Баєса обчислюємо апостеріорні ймовірності $P(Z^j / O)$ належності об'єкта до кожного з описів Z^j :

$$P(Z^j / O) = \frac{P(O / Z^j) \cdot P(Z^j)}{\sum_{j=1}^J P(O / Z^j) \cdot P(Z^j)}, \quad (2.8)$$

де $P(O / Z^j)$ – апіорна ймовірність належності об'єкта до еталону Z^j , $P(Z^j)$ – апіорна ймовірність появи еталону Z^j (вважаємо $P(Z^j) = 1 / J$). Дискретна природа задачі призводить до обчислення набору ймовірностей узагальненого гіпергеометричного розподілу [70]:

$$P(O / Z^j) = \prod_{i=1}^k C_{h_i[Z^j]}^{h_i[O]} / C_H^s, \quad (2.9)$$

де $s = \sum_{i=1}^k h_i [O]$ – потужність множини елементів опису об’єкту.

Формула (2.9) має сенс лише для невід’ємних цілих значень параметрів, з умовами

$$h_i [O] \leq h_i [Z^j], \quad s \leq H, \quad i = 1, \dots, k, \quad j = 1, \dots, J. \quad (2.10)$$

Обмеження (2.10) потребують коригування даних, яке можна здійснити пропорційним змінюванням характеристик об’єкта. Коректування об’єкта при невиконанні обмежень (2.10) пропонуємо здійснювати з округленням результату до цілого числа за допомогою коефіцієнта $\lambda_{кор}^j$ за формулою:

$$h'_i [O] = \lambda_{кор}^j \cdot h_i [O], \quad i = 1, \dots, k, \quad j = 1, \dots, J.$$

Розрахунок треба здійснювати для еталонів Z^j , де порушено умови (2.10). Об’єкт підходить до класу $\nu \in \{1, \dots, J\}$ за правилом

$$\nu = \arg \max_{j=1, \dots, J} P(Z_j / O), \quad (2.11)$$

що є реалізацією оптимізаційної процедури загального вигляду

$$\nu = \arg \operatorname{opt}_{j=1, \dots, J} \Lambda(h[O], h[Z^j]). \quad (2.12)$$

Наведений метод є оптимальним у сенсі найменшої в середньому ймовірності помилки класифікації [45, 73].

2.2.2. Метричні класифікатори

Частотні підходи нерідко називають метричними [68, 72], навіть якщо міра Λ в (2.12) не задовольняє аксіомам метрики. Метричний класифікатор відносить досліджуваний об'єкт до еталону, міра подібності якого виявилася оптимальною. При цьому додатково здійснюють перевірку значущості отриманого мінімуму релевантності відповідно до встановленого порогу [28].

Найпоширенішою є евклідова метрика, яка ефективно використовується у випадках достатньої відмінності еталонів між собою, і часто демонструє нечутливість до варіативності даних у інших ситуаціях.

Останнього часу дослідники методів розпізнавання зображень зосереджують увагу на мірі «розходження Кульбака-Лейблера» (РКЛ) [74-76]. Цю міру для нашої задачі можна інтерпретувати як значення середньої інформації щодо відмінності об'єкту O від еталону Z^j [77]:

$$D_{KL}(O \parallel Z^j) = \sum_{i=1}^k h_i^*[O] \ln \frac{h_i^*[O]}{h_i^*[Z^j]}, \quad (2.13)$$

де $h_i^*[O]$, $h_i^*[Z^j]$ – елементи нормованих описів (2.1).

Сутність міри (2.13) пояснимо теоретичними міркуваннями, що обґрунтовують її застосування для класифікації [77]. Обчислимо апостеріорні ймовірності $P(Z_j / O)$ за формулою (2.8), зокрема, для $j = 1$, $j = 2$ маємо:

$$\frac{P(Z^1 / O)}{P(Z^2 / O)} = \frac{P(O / Z^1) \cdot P(Z^1)}{P(O / Z^2) \cdot P(Z^2)}. \quad (2.14)$$

Після логарифмування (2.14)

$$\ln \frac{P(O/Z^1)}{P(O/Z^2)} = \ln \frac{P(Z^1/O)}{P(Z^2/O)} - \ln \frac{P(Z^1)}{P(Z^2)}. \quad (2.15)$$

Інтерпретація (2.15) при порівнянні двох еталонів [77]:

1) це різниця між логарифмами шансів належності об'єкту до Z^1 та Z^2 до та після спостережень на користь еталону Z^1 , її можна інтерпретувати як інформацію, отриману в результаті спостереження;

2) це логарифм відношення правдоподібності, що трактується як інформація щодо об'єкту O для його розрізнення на користь Z^1 проти Z^2 ;

3) це критерій прийняття рішення, який полягає в обранні еталону, для якого логарифм відношення апостеріорної ймовірності до апіорної є більшим.

Розглянемо тепер дискретну випадкову величину, що приймає значення (2.15), виходячи з нормованих кластерних розподілів. Ліва частина (2.15) є $\ln \frac{h_i^*[O]}{h_i^*[Z^j]}$, $i=1, \dots, k$, що можна розуміти як інформацію щодо об'єкту

O на предмет його відмінності від еталону Z^1 . Вважаємо, що така величина приймає вказані значення з ймовірностями (частотою) $h_i^*[O]$, $i=1, \dots, k$. Обчислення її математичного сподівання приводить до виразу РКЛ (2.13) і інтерпретується як середня інформація щодо розрізнення об'єкту від еталону, або як середня величина втрат інформації при віднесенні O до еталону Z^1 .

Міра (2.13) не є метрикою, оскільки не задовольняє аксіомам симетричності та нерівності трикутника, та є невід'ємною, причому рівність нулю має місце тоді і лише тоді, коли $h_i^*[O] = h_i^*[Z^j]$, $i=1, \dots, k$ [77, 78].

Міра (2.13) відповідає базовому методу математичної статистики, що дає оцінку параметрів розподілу, – методу максимальної правдоподібності, та приводить до вибору розподілу (еталону), що знаходиться на мінімальній відстані (2.13) від невідомого розподілу об'єкта [79].

Узагальненням міри РКЛ є розходження Рен'ї, яке для нашої задачі можна розглядати як інформацію порядку α ($\alpha > 0$, $\alpha \neq 1$) щодо відмінності об'єкту O від еталону Z^j [77]:

$$D_{\alpha}(O \parallel Z^j) = \frac{1}{\alpha - 1} \ln \sum_{i=1}^k \frac{(h_i^*[O])^{\alpha}}{(h_i^*[Z^j])^{\alpha-1}}. \quad (2.16)$$

Розходження Рен'ї також не є метрикою, однак є на сьогоднішній день прикладним інструментарієм для порівняння зображень чи їх описів за рахунок ряду властивостей, що забезпечують зрозуміле з практичної точки зору і ефективно з теоретичного погляду розпізнавання образів [74]. Міра (2.16) приймає невід'ємні значення, а нульове значення досягається лише при $h_i^*[O] = h_i^*[Z^j]$, $i = 1, \dots, k$, і є неспадною функцією аргументу α . При $\alpha \rightarrow 1$ вираз (2.16) приймає вигляд (2.13) і відповідає важливому окремому випадку розходження Рен'ї – мірі РКЛ, тобто, $D_1(O \parallel Z^j) = D_{KL}(O \parallel Z^j)$.

Особливим можна також вважати значення параметра $\alpha = 0,5$, що приводить міру (2.16) до відстані Бхатгачарія [80]:

$$D_{0,5}(O \parallel Z^j) = -2 \ln \sum_{i=1}^k \sqrt{h_i^*[O] \cdot h_i^*[Z^j]}, \quad (2.17)$$

яка, між іншим, є єдиним випадком симетричності розходження Рен'ї відносно обох описів O і Z^j , а також є метрикою. Доведено [80], що оптимальним значенням параметра $\alpha \in [0;1]$ є значення $\alpha = 0,5$. При цьому оптимальним вважається значення параметра, при якому ненормована величина розходження Рен'ї $(\alpha - 1) \cdot D_{\alpha}(O \parallel Z^j)$ приймає найбільше значення, що відповідає нижній границі ймовірності помилки розпізнавання [73].

Таким чином, у випадку значної близькості об'єктів (за їх структурними описами) найкращим для їх розрізнення значенням параметру можна вважати $\alpha = 0,5$, яке у певному сенсі мінімізує величину розходження Рен'ї. Окремий інтерес представляє також значення $\alpha = 2$, яке приводить міру (2.16) до вигляду логарифму математичного сподівання відношення частот $\frac{h_i^*[O]}{h_i^*[Z^j]}$:

$$D_2(O \| Z^j) = \ln \sum_{i=1}^k \frac{(h_i^*[O])^2}{h_i^*[Z^j]}. \quad (2.18)$$

Зауважимо [81], що наведена міра з точністю до монотонного перетворення еквівалентна відстані $\chi^2(O, Z^j)$:

$$D_2(O \| Z^j) = \ln(1 + \chi^2(O, Z^j)), \quad (2.19)$$

де $\chi^2(O, Z^j) = \sum_{i=1}^k \frac{(h_i^*[O] - h_i^*[Z^j])^2}{h_i^*[Z^j]}$ – відстань хі-квадрат, що відповідає

класичному у математичній статистиці критерію χ^2 Пірсона, за яким перевіряється статистична гіпотеза про відсутність відмінностей двох структурних описів.

2.2.3. Порівняльний аналіз прикладних властивостей мір релевантності

Проаналізуємо найбільш вагомі особливості вказаних мір з точки зору практичного застосування моделі оптимізаційного класифікатора (2.12) [13-16]. Евклідова метрика є однією з найбільш вживаних мір подібності [22, 72].

Однак практичний досвід показує, що вона недостатньо чутлива щодо варіативності даних.

Покажемо це на прикладі. Нехай опис об'єкту O і еталонів Z^1, Z^2 подається даними таблиці 2.7.

Таблиця 2.7. Кластерні подання O, Z^1, Z^2

	Кластер		
	M_1	M_2	M_3
O	5	3	2
Z^1	8	8	4
Z^2	12	6	2

Обчислимо евклідову відстань (E) між об'єктом і еталонами, розходження Рен'ї при різних значеннях параметра α ($D_{0,5}(O \parallel Z^j)$, $D_{KL} = D_1(O \parallel Z^j)$, $D_2(O \parallel Z^j)$), а також апостеріорні ймовірності (P) віднесення об'єкта до еталонів (таблиця 2.8).

Таблиця 2.8. Значення мір релевантності за даними таблиці 2.7

	E	$D_{0,5}(O \parallel Z^j)$	D_{KL}	$D_2(O \parallel Z^j)$	P
Z^1	0,141	0,013	0,025	0,049	0,543
Z^2	0,141	0,022	0,047	0,110	0,457

Як бачимо із таблиці 2.8, в той час як евклідова відстань між об'єктом і еталонами однакова, решта мір свідчать про перевагу еталону Z^1 . Цей приклад в цілому демонструє переважаючу чутливість статистичних мір $D_{0,5}(O \parallel Z^j)$, D_{KL} , $D_2(O \parallel Z^j)$ та P у порівнянні із евклідовою метрикою та підкреслює важливість вибору адекватної міри релевантності.

Аналіз особливостей практичного застосування цих двох мір релевантності виконаємо, спираючись на приклад опису 4-х еталонів в поданні із 10-ти кластерів за даними роботи [13, 22].

Таблиця 2.9. Кластерні подання бази $Z = \{Z^j\}_{j=1}^4$

Еталон	Кластер									
	M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8	M_9	M_{10}
Z^1	7	9	9	9	14	15	6	11	11	9
Z^2	8	11	8	16	3	14	13	9	11	7
Z^3	15	7	3	13	9	18	7	8	11	9
Z^4	10	8	8	13	13	12	7	11	11	7

Результати обчислення мір релевантності об'єктів, в якості яких взято безпосередньо еталони, наведено в таблиці (2.10).

Таблиця 2.10. Значення мір релевантності за даними таблиці (2.9)

Еталон	Об'єкт Z^1				
	P	$D_{0,5}(O \parallel Z^j)$	D_{KL}	$D_2(O \parallel Z^j)$	E
Z^1	0	0	0	0	0,99
Z^2	0,15	0,069	0,156	0,398	4,9E-07
Z^3	0,13	0,046	0,095	0,204	8,0E-08
Z^4	0,01	0,010	0,020	0,038	6,7E-08
Еталон	Об'єкт Z^2				
	P	$D_{0,5}(O \parallel Z^j)$	D_{KL}	$D_2(O \parallel Z^j)$	E
Z^1	0,15	0,069	0,124	0,213	3,0E-10
Z^2	0	0	0	0	0,99
Z^3	0,14	0,059	0,116	0,225	2,0E-08
Z^4	0,13	0,052	0,090	0,146	1,4E-09

Еталон	Об'єкт Z^3				
	P	$D_{0,5}(O \parallel Z^j)$	D_{KL}	$D_2(O \parallel Z^j)$	E
Z^1	0,13	0,046	0,089	0,171	2,6E-09
Z^2	0,14	0,059	0,120	0,246	3,5E-08
Z^3	0	0	0	0	0,99
Z^4	0,11	0,031	0,059	0,108	1,0E-10
Еталон	Об'єкт Z^4				
	P	$D_{0,5}(O \parallel Z^j)$	D_{KL}	$D_2(O \parallel Z^j)$	E
Z^1	0,01	0,010	0,021	0,045	1,5E-07
Z^2	0,13	0,052	0,120	0,327	8,2E-07
Z^3	0,11	0,031	0,065	0,144	5,0E-07
Z^4	0	0	0	0	0,99

Порівняння значень показників таблиці 2.10 свідчить про позитивний результат для усіх мір щодо правильного прийняття рішення в межах бази еталонів, високу точність розпізнавання баєсовського класифікатора, що підтверджується суттєвими відмінностями розрахованих значень апостеріорних ймовірностей, а також приблизно рівну чутливість до відхилення значень евклідової відстані та розходження Рен'ї.

Аналіз теоретичних міркувань та кількісних розрахунків дає підстави стверджувати про ряд недоліків при застосуванні баєсовського класифікатора, серед яких – необхідність здійснювати громіздкі обчислення над ненормованими даними, деяка втрата точності через потребу коригування описів, а також виконання обмежень згідно умов (2.10).

У той самий час розходження Рен'ї, хоча і не є традиційною метрикою, має вагомі практичні переваги, що свідчать на користь цієї міри, до яких відноситься відсутність обмеження щодо типу розподілу спостережень, який передбачається невідомим. При цьому вимога відсутності нульових значень у структурних описах еталонів (що впливає з (2.13)) може бути задоволена за рахунок коригування опису об'єкта або укрупненням кластерного розподілу.

Відмітимо окремо важливий етап процесу класифікації, пов'язаний із пороговою верифікацією результатів оптимізаційного класифікатора (2.12).

Різна природа аналізованих мір релевантності обумовлює відмінні підходи до формування відповідних порогів, які визначаються рівнем розрізнення в конкретній базі еталонів і рівнем перешкод [13-16, 22]. Баєсовський класифікатор потребує встановлення нижнього рівня для значення апостеріорної ймовірності. Класифікатори на основі мір подібності передбачають неперевищення деякого порога для мінімуму відстані. В свою чергу, розходження Рен'ї припускає (через свою статистичну сутність) використання апарату математичної статистики в розрізі перевірки статистичних гіпотез. Так, наприклад, в [80] пропонується встановлення порогу $\rho = \frac{1}{2s} \chi_{k-1,1-a}^2$ для міри РКЛ, призначеного для відмови від віднесення досліджуваного об'єкта до жодного з еталонів $Z^j, j=1, \dots, J$, якщо виконується умова $D_{KL}(O \| Z^j) \geq \rho$, де a – задана ймовірність помилки першого роду (ймовірність відхилення гіпотези про віднесення об'єкта до еталону за умови, що вона виявилася справедливою), $\chi_{k-1,1-a}^2$ – критична точка розподілу χ^2 відповідно до числа ступенів свободи $k-1$ та рівня значущості $1-a$. В умовах наведеного прикладу за даними таблиці 2.9 маємо $s=100, k=10$, звідки при $a=0,05$ одержуємо $\rho=0,017$. Як бачимо тепер з таблиці 2.10, всі розраховані значення РКЛ еталонів вище отриманого порогу, що, зокрема, свідчить про значущу відмінність еталонів між собою.

Зауважимо також, що перевірка гіпотези про відсутність відмінностей між еталонами за даними таблиці 2.9 на підставі критерію χ^2 Пірсона [79], проведена на підставі результатів розрахунків розходження Рен'ї $D_2(O \| Z^j)$ при $\alpha=2$, показала, що критична точка розподілу $\chi_{k-1,a}^2=16,9$ (при рівні значущості $a=0,05$ і при числі ступенів свободи $k-1=9$) не перевищує найменші з відповідних ненульових спостережених значень критерію (за кожним з досліджуваних об'єктів окремо), з чого випливає неможливість відкинути гіпотезу про відсутність відмінностей (наприклад, між еталоном

Z^1 як об'єктом і еталоном Z^4 , між еталоном Z^2 як об'єктом і еталоном Z^4 тощо), що може свідчити про недостатню чутливість розходження Рен'ї при $\alpha = 2$ щодо відмінності досліджуваних об'єктів.

Висновки

1. Застосування статистичних підходів при обчисленні мір релевантності описів дає можливість не тільки покращити показники розрізнення за рахунок більш значущої чутливості статистичних мір, а також і здійснити верифікацією результатів оптимізаційного класифікатора.

2. Статистичні міри мають детальне математичне обґрунтування та в процесі класифікації оптимізують задані критерії, вважаючи ймовірнісний розподіл об'єкта невідомим. Статистичне подання відображає більш загальні та суттєві властивості об'єкта та еталона, що зосереджені в їх образах у вигляді розподілів значень класифікаційних ознак.

3. За результатом дослідження робимо висновок про доцільність використання розходження Рен'ї (при $\alpha = 1$ і $\alpha = 0,5$) для структурної класифікації, оскільки ця міра є чутливою до різноманіття кластерних подань та має універсальне застосування для довільного типу даних.

3. Визначення релевантності зображень шляхом застосування статистичного аналізу для фрагментного подання їх описів

3.1 Формалізація побудови системи статистичних розподілів

Визначимо опис розпізнаваного візуального об'єкту у вигляді скінченної множини $Z = \{z_v\}_{v=1}^s$, $z_v \in B^n$ із числа s бінарних дескрипторів КТ, сформованих, наприклад, детекторами ORB, BRISK [38, 39], B^n – множина усіх бінарних векторів розмірності n .

Традиційний метод для обчислення релевантності описів $O_1, O_2 \subset B^n$ полягає у підрахунку числа r «еквівалентних» елементів одного з описів, що вважається еталоном, у складі іншого опису [28]

$$r = \text{card}\{x_1 \in O_1 \mid x_1 \sim x_2, x_2 \in O_2\}. \quad (3.1)$$

Еквівалентність дескрипторів $x_1 \sim x_2$ в (3.1) трактується як близькість у векторному просторі

$$x_1 \sim x_2 : \rho(x_1, x_2) \leq \delta, \quad (3.2)$$

де ρ – метрика у просторі B^n , δ – поріг для значення метрики, що задає околицю, в межах якої елементи x_1, x_2 вважаються еквівалентними.

Як правило, прийняття рішення ґрунтується на нормуванні значення (3.1):

$$r^* = r / m_1, \quad (3.3)$$

де $m_1 = \text{card}\{O_1\}$ – чисельність елементів еталонної множини O_1 .

Для дескрипторів КТ із простору B^n у виразі (3.2) можна використати нескладну для обчислень метрику Хемінга. Більш проблематичні за обсягом обчислень підходи пов'язані з визначенням мір подібності чи відстані у просторі описів як множин КТ [28].

Для кожного $z_v \in Z$ виконаємо розбиття на m непересічних фрагментів однакового розміру: $z_v = z_v^1 \& z_v^2 \& \dots \& z_v^m$, де z_v^k – значення k -го фрагменту дескриптора z_v . Кожний із дескрипторів тепер складається із m однотипних елементів, послідовність яких можна записати набором цілих чисел відомого діапазону. Перетворений опис Z загалом тепер розглядаємо як матрицю із s рядків по m елементів у рядку [12-18].

Побудуємо відображення $\Omega: B^n \rightarrow Q$ із множини бінарних векторів – дескрипторів КТ у множину Q цілочисельних векторів – статистичних розподілів їх фрагментів меншої потужності $w < n$. Відображення Ω повинне забезпечувати можливість ідентифікувати та розрізняти візуальні об'єкти у просторі розподілів їх значень.

Визначимо розподіл $q \in Q$ як вектор $q = \{q_1, \dots, q_w\}$ цілих чисел, де q_i – це кількість елементів фіксованого фрагмента серед елементів множини Z , що приймають значення i , $i = \overline{1, w}$. Тоді $\sum_i q_i = s$, тобто сума значень компонентів у розподілі q дорівнює загальному об'єму s аналізованих дескрипторів опису Z .

Для кожного із m фрагментів на підставі аналізу множини Z побудуємо розподіл $q = \{q_1, \dots, q_w\}$, причому величина w цілком визначається діапазоном значень даних для фрагмента, що у бінарному поданні прямо залежить від розміру фрагмента. Наприклад, для дескриптора BRISK при розбитті на байти для $n = 512$ маємо $m = 64$, $w = 256$. Повністю множина даних Z під дією відображення Ω буде описана матрицею розподілів $Q = \{q^v = (q_1, q_2, \dots, q_w)^v\}_{v=1}^m$ для побудованої системи із m фрагментів.

Матриця Q відображає статистичні властивості опису Z у вигляді набору розподілів значень його сегментів. Статистичні розподіли даних відповідають фундаментальній концепції машинного навчання, так як здійснюють узагальнення знань [54-58].

У той же час у побудованому просторі розподілів для фрагментів даних у процесі побудови статистичних характеристик здійснюється оцінювання параметрів запропонованої моделі на підставі наявного опису візуального об'єкту. Результат розпізнавання об'єкту визначається шляхом оптимізації значень обчисленого ступеня релевантності як найбільш близький еталон відповідного класу.

3.2. Оцінювання значущості прийняття класифікаційних рішень для моделі блочного подання

Розглянемо відображення $Z \rightarrow Q$ із множини бінарних векторів – дескрипторів КТ у множину Q їх блочних розподілів, яке дає можливість проводити ідентифікацію та розрізнення візуальних об'єктів на підставі об'єктів меншої розмірності [14, 15]. Розподіл відповідно до фіксованого фрагменту із номером $j, j=1, \dots, m$, подамо у вигляді вектора цілих чисел $q_j = (q_1, \dots, q_w)_j$, де $q_{ij}, i=1, \dots, w$, – частота появи i -го виду фрагменту, розташованого на j -му місці ланцюжка серед усіх дескрипторів опису, кількість яких дорівнює $s = \sum_{i=1}^w q_i$ для будь-якого номера j .

Значення $p_i = \frac{q_i}{s}, i=1, \dots, w$, є відносними частотами (оцінками ймовірності). Іноді для них використовують назву «апостеріорні ймовірності». При цьому $w = 2^k, \sum_{i=1}^w p_i = 1$. Повністю множина Q описується системою розподілів

$$Q = \left\{ q_j = (q_1, \dots, q_w)_j \right\}_{j=1}^m \quad (3.4)$$

для кожного із m фрагментів, отже, може бути представлена матрицею із m рядків та w стовпців. Для нормованого випадку маємо систему розподілів

$$Q_{\text{норм}} = \left\{ p_j = (p_1, \dots, p_w)_j \right\}_{j=1}^m. \quad (3.5)$$

Наприклад, у випадку використання дескриптора BRISK при розбитті на байти для $n = 512$ маємо $m = 64$, $w = 256$, отже, матриця Q ($Q_{\text{норм}}$) містить відносні частоти появи всіх допустимих значень фрагментів розміром у байт для наявної множини Z .

Основна ідея дослідження полягає у здійсненні аналізу релевантності об'єктів шляхом зіставлення сформованих значень елементів матриць $Q^l = \{q_{ij}^{(l)}\}$, $i = 1, \dots, w$, $j = 1, \dots, m$, $l = 1, 2$, (l – номер об'єкту) вигляду (3.4), (3.5), що відображають статистичні властивості описів. Зосередимо дослідження на попарному порівнянні відповідних рядків матриць Q^l .

Першочерговою видається вирішення задачі порівняння двох розподілів, яку можна здійснити співставленням гістограм через застосування модифікованого критерію хі-квадрат, призначеного для випадку емпіричних розподілів [82]. Розрахунок емпіричного значення критерію для фрагментів обох описів з номером j здійснимо за формулою

$$\chi_{\text{емп}, j}^2 = \sum_{i=1}^w \frac{(q_{i,j}^{(1)} - q_{i,j}^{(2)})^2}{q_{i,j}^{(1)} + q_{i,j}^{(2)}}. \quad (3.6)$$

Критичне значення критерію $\chi_{кр,j}^2(a,s')$ знаходиться за встановленим рівнем значущості a і ступенем свободи s' , що не перевищує $s-1$ (s розглядається як кількість спостережень) і розраховується лише для тих типів ланок, які входять до відповідного фрагменту з опису хоча б одного з об'єктів. Область прийняття нульової гіпотези про однорідність аналізованих гістограм визначається нерівністю $\chi_{емп,j}^2 < \chi_{кр,j}^2(a,s')$, а область прийняття альтернативної гіпотези щодо значущої відмінності між ними на рівні значущості a є правосторонньою критичною областю, що визначається нерівністю $\chi_{емп,j}^2 > \chi_{кр,j}^2(a,s')$. Наступний крок передбачає підрахунок долі значущо відмінних розподілів, на підставі чого робиться висновок про рівень релевантності об'єктів. Природнім видається очікування підвищення точності розпізнавання при збільшенні довжини k фрагменту у ланцюжковому поданні.

Іншим способом зіставлення розподілів є застосування в якості міри релевантності α -розходження Рен'ї [9, 77], яке можна розглядати як інформацію порядку α ($\alpha > 0, \alpha \neq 1$) щодо їх відмінності. Обчислення розходження Рен'ї $D_{\alpha j}(p_j^{(1)} \| p_j^{(2)})$, $j=1, \dots, m$ щодо відмінності об'єкту з розподілом $p_j^{(1)} = (p_1, \dots, p_w)_j^{(1)}$, від об'єкту з розподілом $p_j^{(2)} = (p_1, \dots, p_w)_j^{(2)}$ (для кожного фіксованого $j=1, \dots, m$), здійснюємо, виходячи з рядків матриць $Q_{норм}^l = \{p_{ij}^{(l)}\}$, $i=1, \dots, w, j=1, \dots, m, l=1, 2$, що саме і мають вигляд $p_j^{(l)} = (p_1, \dots, p_w)_j^{(l)}$, $j=1, \dots, m, l=1, 2$:

$$D_{\alpha j}(p_j^{(1)} \| p_j^{(2)}) = \frac{1}{\alpha - 1} \ln \sum_{i=1}^w \frac{(p_{ij}^{(1)})^\alpha}{(p_{ij}^{(2)})^{\alpha-1}}. \quad (3.7)$$

Міра (3.7) є неспадною функцією аргументу α , приймає невід'ємні значення, а нульове значення досягається лише при $p_j^1 = p_j^2$. Пропонується застосування міри (3.7) при значенні параметра $\alpha = 0,5$:

$$D_{0,5;j}(p_j^{(1)} \parallel p_j^{(2)}) = -2 \ln \sum_{i=1}^w \sqrt{p_{ij}^{(1)} \cdot p_{ij}^{(2)}}, \quad j = 1, \dots, m. \quad (3.8)$$

Зауважимо, що міра (3.8) є метрикою, а також єдиним випадком симетричності розходження Рен'ї відносно аргументів. Доведено [73, 80], що $\alpha = 0,5$ є оптимальним у тому сенсі, що величина $(\alpha - 1) \cdot D_\alpha(O \parallel Z^j)$ при $\alpha = 0,5$ приймає найбільше значення, яке відповідає нижній границі ймовірності помилки розпізнавання. З цієї точки зору значення параметру $\alpha = 0,5$ можна вважати найкращим для розрізнення структурних описів.

З метою всебічного оцінювання доцільності введення ланцюжкової структури для опису об'єкта і аналізу його статистичного подання як ефективного підходу для розрізнення об'єктів пропонуємо перехід від блочного подання цих описів до опису із кодуванням структур фрагмента із заміною розподілу фрагмента числом, яке дорівнює середньому кодових значень. Наприклад, у випадку 2-бітового розбиття дескриптора, одержаного детектором BRISK, кодуємо числами від 1 до 4 послідовність 00, 01, 10, 11, що включає можливі значення фрагмента. Згідно із статистичним поданням виду (2), тепер маємо в межах одного фрагмента з номером j розподіл вигляду $p_j = (p_1, p_2, p_3, p_4)_j$. Відповідно до кодових значень 1, 2, 3, 4, математичне сподівання (середнє значення) розподілу дорівнює:

$$M_j = \sum_{i=1}^4 i p_{ij}, \quad j = 1, \dots, 256. \quad (3.9)$$

Тоді статистичне подання (3.5), що складається з 256-ти рядків та 4-х стовпців, перетворюється на послідовність із 256-ти елементів, кожний з яких є усередненням фрагмента:

$$M = (M_1, \dots, M_{256}). \quad (3.10)$$

У загальному випадку вирази (3.9), (3.10) для об'єкта з номером $l=1,2$, набувають узагальненого вигляду

$$M_j^{(l)} = \sum_{i=1}^w ip_{ij}^{(l)}, \quad j = 1, \dots, m, \quad (3.11)$$

$$M^{(l)} = (M_1, \dots, M_m)^{(l)}. \quad (3.12)$$

Порівняння двох описів здійснюємо через співставлення різними способами двох кортежів виду (3.12). Наприклад, порівняння послідовностей $M^{(1)}$, $M^{(2)}$ як двох зв'язних вибірок довільного розподілу можливе із застосуванням непараметричних критеріїв відмінності, наприклад, критерію знаків [83], який на встановленому рівні значущості показує, наскільки односпрямованими за знаком є різниці

$$M_i^{(2)} - M_i^{(1)}, \quad i = 1, \dots, w. \quad (3.13)$$

Згідно із критерієм знаків розраховуємо окремо суми додатних і від'ємних різниць (3.13), більша з них призначається типовим зсувом, менша – нетиповим. За таблицями [83] значення типового зсуву зіставляється (на певному рівні значущості) з критичним значенням, неперевищення якого величиною нетипового зсуву свідчить про значущу відмінність вибірок.

Іншим способом зіставлення послідовностей $M^{(1)}$, $M^{(2)}$ може бути попарне порівняння значень $M_j^{(1)}$ та $M_j^{(2)}$, $j=1, \dots, m$, послідовно для розподілу кожного фрагмента як двох середніх для довільно розподілених сукупностей за допомогою z -критерію, що є коректним лише у випадку достатньо великих вибірок [83] і узгоджується з вимогою достатньо великого числа s . Розрахунок емпіричного значення критерію здійснимо як

$$z_{em,j} = \frac{|M_j^{(1)} - M_j^{(2)}|}{\sqrt{(D_j^{(1)} + D_j^{(2)}) / s}}, \quad j=1, \dots, m, \quad (3.14)$$

де $D_j^{(l)} = \sum_{i=1}^w p_{ij}^{(l)} (i - M_j^{(l)})^2$, $l=1,2$, – дисперсія вибірки, заданої розподілом $p_j^{(l)} = (p_1, \dots, p_w)_j^{(l)}$, $l=1,2$, із відповідними кодовими значеннями $i=1, \dots, w$.

Критичне значення критерію $z_{кр,j}(a)$ знаходиться за рівнем значущості a із таблиці значень функції Лапласа $\Phi(x)$: $\Phi(z_{кр,j}(a)) = (1 - 2a) / 2$.

Область прийняття нульової гіпотези про рівність усереднених кодових значень в межах j -го фрагмента визначається нерівністю $z_{em,j} < z_{кр,j}(a)$, область прийняття альтернативної гіпотези визначається нерівністю $z_{em,j} > z_{кр,j}(a)$. Наступний крок передбачає підрахунок долі таких розподілів фрагментів, відповідні номери яких для обох об'єктів мають значущо відмінні усереднені кодові значення, на підставі чого робиться висновок про рівень релевантності об'єктів. Природним видається очікування достатньо великого значення такої долі для об'єктів, що дійсно є різними, а також його підвищення при збільшенні фрагмента. Експериментальна частина роботи включає результати відповідного тестування.

Запропоновані підходи протестовано на прикладі зображень ікон розміром 400x540 із використанням засобів бібліотеки Open CV [12-18, 86]. Застосовано дескриптори BRISK розмірністю $n=512$ та числом 700. Приклад розподілу значень описів для пар бітів наведено у таблиці 3.1.

Таблиця 3.1. Приклад розподілів для 2-бітового розбиття

Фрагмент	Ланки для ікони 1				Ланки для ікони 2			
	00	01	10	11	00	01	10	11
1	468	21	22	189	459	47	25	169
2	213	101	236	150	175	102	263	160
...
256	357	90	58	195	351	62	64	223

Як бачимо із табл. 3.1, розподіли фрагментів з однаковими номерами для обох об'єктів досить суттєво відрізняються, що має бути підтверджено відповідними розрахунками.

Проведемо дослідження однорідності розподілів попарно між фрагментами для 2-х, 4-х та 8-ми бітового розбиттів перевіркою статистичної гіпотези про їх однорідність за критерієм χ^2 -квадрат (3.6). У таблиці 3.2 наведено результати застосування критерію.

Таблиця 3.2. Число неоднорідних розподілів за критерієм χ^2 -квадрат

	Число бітів блочного розбиття		
	2	4	8
Число фрагментів із значущою неоднорідністю розподілів (на рівні 0,05)	167 із 256 (64,45 %)	109 із 128 (85,16 %)	64 із 64 (100%)

Як бачимо, неоднорідність розподілів стає більш вираженою при зростанні розміру фрагмента.

Підвищення рівня розрізнення об'єктів за рахунок подовження фрагментів у ланцюжковій структурі демонструє також розрахунок α -розходження Рен'ї (при $\alpha=0,5$) між розподілами відповідно до виразу (5). У таблиці 3.3 представлено діапазони значень міри релевантності, звідки бачимо суттєве зростання показника при збільшенні кількості бітів у фрагменті.

Таблиця 3.3. Значення α -розходження Рен'ї ($\alpha=0,5$) в залежності від числа бітів у фрагментах

Значення α -розходження Рен'ї	Число бітів блочного розбиття		
	2	4	8
Мінімальне	0,00008	0,00397	0,05111
Максимальне	0,02112	0,03248	0,18575

Дані табл. 3.3 демонструють суттєве підвищення якості розрізнення двох різних об'єктів. Так, наприклад, найменше з можливих значень міри подібності при 8-бітовому типі розбиття на фрагменти (0,05111) перевищує найбільше з можливих значень міри подібності при 4-бітовому типі розбиття (0,03248).

Продовжимо статистичний аналіз рівня відмінностей образів, виходячи із нормованих розподілів, і обчислимо для кожного розподілу фрагмента математичне сподівання, знижуючи тим самим розмірність структури опису. Наприклад, у випадку 2-бітового розбиття маємо статистичні розподіли для першого фрагмента 1-го і 2-го об'єктів відповідно (0,669; 0,030; 0,031; 0,270) і (0,656; 0,067; 0,036; 0,241) із математичними сподіваннями $M_1^{(1)} = 1,903$, $M_1^{(2)} = 1,862$. При цьому статистичні подання (3.5) перетворюються на послідовності із 256-ти елементів:

$$M^{(1)} = (1,903; 2,461; \dots; 2,130), M^{(2)} = (1,862; 2,583; \dots; 2,227). \quad (3.15)$$

Дослідження на однорідність отриманих у такий спосіб узагальнених описів об'єктів на основі критерію знаків приводить до висновку про їх значущу відмінність (на рівні 0,05) при різних довжинах фрагментів. У таблиці 3.4 наведено результати застосування критерію.

Таблиця 3.4. Результати застосування критерію знаків

Кількість бітів у фрагменті	Кількість типових зсувів	Кількість нетипових зсувів	Критичне значення (рівень значущості 0,05)	Висновок щодо значущості
2	197	59	87	Значуща (59 < 87)
4	100	27	41	Значуща (27 < 41)
8	48	16	17	Значуща (16 < 17)

Аналіз таблиці 3.4 показує, що для різної кількості бітів у фрагменті спостерігається наявність значущої статистичної відмінності між описами двох об'єктів, візуально досить схожих між собою, що в черговий раз підтверджує доцільність ідеї щодо блочного подання даних дескрипторів ключових точок їх зображення. Зауважимо, що, як видно з таблиці 3.4, навіть зменшення довжин послідовностей вигляду (3.12) при відповідному збільшенні довжини фрагментів демонструє значущу розрізненість цих послідовностей, що, в свою чергу, є підтвердженням розрізненості досліджуваних об'єктів.

У межах останнього підходу з використанням послідовностей вигляду (3.12), складених із математичних сподівань кодів фрагментів, виконано попарне порівняння $M_j^{(1)}$ і $M_j^{(2)}$ послідовно для $j, j = 1, \dots, m$, із використанням z-критерію. Розрахунок емпіричного значення критерію для

фрагмента з номером $j=1, \dots, m$ здійснювався за формулою (3.14), критичне значення критерію на рівні значущості 0,05 становить 1,64 [14]. Наприклад, для послідовностей (3.15) значення $M_1^{(1)} = 1,903$, $M_1^{(2)} = 1,862$ значущо не відрізняються, а значення $M_2^{(1)} = 2,461$, $M_2^{(2)} = 2,583$ є значущо відмінними (на рівні 0,05). У таблиці 3.5 зведено основні результати застосування критерію.

Таблиця 3.5. Кількість значущих відмінностей за z-критерієм

	Число бітів блочного розбиття		
	2	4	8
Число фрагментів із значущими відмінностями (на рівні 0,05)	148 із 256 (57,81 %)	84 із 128 (65,63 %)	44 із 64 (68,75 %)

Як видно з таблиці 3.5, розрахунки за z-критерієм є ще одним свідченням щодо встановлення значимого розрізнення об'єктів, які реально є різними, а якість розрізнення зростає при збільшенні параметра кількості бітів у фрагменті.

3.3. Обчислення релевантності та класифікація за значеннями розподілів

Визначення релевантності полягає у встановленні міри для відмінності наявних значень матриць $Q = \{q_{i,j}\}$, $i=\overline{1,w}$, $j=\overline{1,m}$, отриманих для описів окремих об'єктів.

Визначимо релевантність r для описів a та b на підставі зіставлення множини їх розподілів щодо введеної системи фрагментів як манхетенську відстань між матрицями $Q(a)$, $Q(b)$:

$$r[Q(a), q(b)] = \sum_{i=1}^w \sum_{j=1}^m |q_{i,j}(a) - q_{i,j}(b)|. \quad (3.16)$$

Для оцінювання вагомості обчисленого значення релевантності (3.16) у межах визначеного її діапазону запровадимо також нормовану відстань

$$r^* = r / r_{\max}, \quad (3.17)$$

де r_{\max} – максимальна відстань між двома розподілами при фіксованому числі s точок опису, розміру n дескриптора і кількості бітів у розподілі.

Наприклад, для одно-бітового розподілу при $s = 700$, $n = 512$ значення $r_{\max} = 1400 \times 512$, для двох-бітового – $r_{\max} = 1400 \times 256$, для 8-ми бітового – $r_{\max} = 1400 \times 64$.

Таким чином, запропонований метод розпізнавання об'єкту можна викласти у вигляді послідовності етапів оброблення:

- 1) формування опису об'єкта як множини дескрипторів КТ;
- 2) виконання розбиття кожного із дескрипторів у вигляді встановленої системи фрагментів;
- 2) побудова статистичних розподілів за множиною фрагментів опису;
- 3) обчислення значень релевантності (3.16) у просторі розподілів об'єкта та еталонів;
- 4) визначення класу з оптимальним значенням релевантності.

Результати комп'ютерного моделювання методів розпізнавання на підґрунті обчислення відстані між розподілами, що побудовані для даних структурних описів, наведені у підрозділі 4.1.

3.4. Просторово-статистичний аналіз та оброблення ансамблю фрагментів

Методи розпізнавання візуальних об'єктів за описом множини КТ отримали практичне застосування із-за таких визначних властивостей, як

інваріантність значень дескрипторів КТ до геометричних перетворень об'єктів, можливість розпізнавання в умовах часткової видимості, стійкість до завад.

До структурних ознак, окрім безпосередньо дескрипторів КТ та їх координат, відносять також атрибути проєкцій, відгуки спеціалізованих локальних фільтрів та ін. Актуальною залишається задача витягнення нової та більш глибокої за властивостями розрізнення об'єктів інформації із наявного обсягу даних, що містить множина дескрипторів. За результатами аналізу або трансформації опису цілеспрямовано може бути сформована нова система ознак із кращими властивостями щодо результативності розпізнавання чи швидкодії оброблення. Головною вимогою для застосування таких методів аналізу чи оброблення даних є забезпечення інваріантності значень побудованих ознак щодо довільної послідовності слідування дескрипторів у описі.

При зіставленні множин дескрипторів у процесі розпізнавання доцільним є застосування факту просторового розміщення компонентів даних опису аналогічно схемі аналізу значень просторових фрагментів сигналу [26]. Врахування просторових властивостей даних дає можливість обробляти звужений піддіапазон значень, що значно спрощує обчислення. Особливі переваги такий аналіз має у зв'язку з можливістю зниження розмірності простору сигналів, наприклад, аналогічно формуванню одновимірних проєкцій для двовимірних об'єктів. Очевидно, що якщо два сигнали повністю ідентичні, то їх окремі фрагменти також еквівалентні, але зворотне твердження не зовсім вірне [28]. За рахунок впровадження просторового аналізу дескрипторів є можливість створити нову структуру даних зі своїми властивостями і діапазонами значень. Якщо обчислювальний ресурс чи обмеження на обсяг пам'яті не дають можливості побудувати багатовимірну гістограму даних, то можна побудувати роздільні гістограми за окремими ознаками. Система може працювати за кількома спрощеними

незалежними каналами, віддаючи перевагу найбільш вірогідному із них або приймаючи рішення за підмножиною каналів [28].

Основним числовим критерієм при розпізнаванні є значення релевантності описів об'єктів, яка відображає статистичну схожість складу аналізованого структурного опису в межах заданої системи еталонів.

Задачами дослідження є створення моделей для обчислення релевантності структурних описів, ефективних за швидкістю оброблення даних, аналіз властивостей цих моделей та експериментальне оцінювання результативності запропонованих підходів за наслідками оброблення конкретних зображень.

Бітова природа дескрипторів КТ у просторі B^n бінарних векторів (n – ступінь двійки) дає можливість запровадити розгляд та оброблення дескриптора як кортежу послідовних елементів (наприклад, байтів), діапазон значень яких відомий. Це дає змогу розглядати структурний опис (множина бінарних векторів) у новоствореному просторі даних та здійснювати його статистичний аналіз з урахуванням властивостей його внутрішнього складу.

Для фіксованого числа n бітів опис об'єкту подаємо бінарною матрицею $D = \{ \{ d_{i,j} \}_{i=1}^s \}_{j=1}^n$ даних, сформовану довільною послідовністю s дескрипторів. Для спрощення аналізу чи оброблення вважаємо значення s однаковим для всіх еталонів, цього нескладно досягти відбором необхідної кількості дескрипторів.

Кожний із рядків матриці D , що є дескриптором КТ, подамо у виді m непересічних фрагментів (блоків), що послідовно слідуєть один за одним. Отримаємо новий простір даних, перевагою якого є можливість аналізувати внутрішню структуру та статистичні властивості опису, на основі чого його можна буде відносити до одного із еталонних класів. Застосуємо мережу блоків як основу для прийняття рішення про клас об'єкту. Блоки дескриптора відіграють роль системи ознак. Такий спосіб формування ознак надає додаткову гнучкість при практичній реалізації [12, 13].

Після розбиття $\forall z_v \in Z$ на m непересічних блоків для кожного дескриптора z_v отримуємо фіксовану структуру $z_v = z_v^1 \& z_v^2 \& \dots \& z_v^m$ елементів, де z_v^k – k -й блок дескриптора z_v . Кожний із дескрипторів тепер складається із m однотипних бінарних елементів, значення яких у послідовності можна записати кортежем цілих чисел відомого діапазону. У результаті опис Z набуває виду матриці із s рядків по m елементів (число блоків) у рядку (таблиця 3.6).

Таблиця 3.6. Блочне подання структури даних D

Номери дескрипторів	Номери блоків				
	1	...	k	...	m
1	z_1^1	...	z_1^k	...	z_1^m
...
s	z_s^1	...	z_s^k	...	z_s^m

Зважаючи на впроваджену рівноцінність діапазонів для значень кожного із фрагментів та відображення сукупністю дескрипторів властивостей аналізованого об'єкту, побудуємо розподіли даних на підставі аналізу матриці D . Для цього задамо відображення

$$\Omega: Z \rightarrow Q, Z \subset B^n$$

із множини бінарних векторів у множину Q цілочисельних векторів потужності $w < n$, що містять статистичні розподіли значень їх фрагментів. Відображення Ω призначене забезпечити ідентифікацію та розрізнення візуальних об'єктів у просторі значень розподілів їх фрагментів.

Визначимо розподіл (рейтинг зустрічальності) $q \in Q$ як вектор $q = \{q_1, \dots, q_w\}$ цілих чисел, де q_i – це число значень для k -го фрагмента, що дорівнюють i , $i = \overline{1, w}$, серед відповідних фрагментів усієї множини Z

$$q_i = \text{card}\{z_v^k \in z_v, z_v \in Z \mid z_v^k = i\}. \quad (3.18)$$

Тоді виконується $\sum_i q_i = s$, тобто сума значень компонентів для розподілу q дорівнює загальному обсягу s аналізованих дескрипторів опису Z . Зауважимо, що w означає число ланок розподілу, а самі значення даних знаходяться у діапазоні $0, \dots, w-1$.

Для кожного із m фрагментів на підставі аналізу побудуємо розподіл $q = \{q_1, \dots, q_w\}$, причому величина w визначена діапазоном значень даних для фрагмента, що у бінарному поданні прямо залежить від його розміру. Наприклад, для дескриптора BRISK при розбитті на байти для $n = 512$ маємо $m = 64$, $w = 256$. Повністю множина даних Z під дією відображення Ω буде описана матрицею $Q = \{\{q_{i,v}\}_{i=1}^w\}_{v=1}^m$, що містить розподіли для побудованої системи фрагментів.

Матриця Q відтворює статистичні властивості опису Z у вигляді кортежу розподілів значень фрагментів його складових. Функціонально статистичні розподіли даних відповідають фундаментальній концепції машинного навчання, здійснюючи узагальнення знань [55-60].

Зауважимо, що обсяг даних матриці Q для незначних розмірів m фрагмента значно менше у порівнянні з матрицею D . Це сприяє суттєвому зменшенню об'єму обчислень при визначенні ступеня релевантності аналізованих описів із еталонними. Однак зі зростом m ситуація дещо змінюється, при $m=8$ об'єми даних приблизно співпадають, а для $m > 8$ обсяг даних за розподілами починає перевищувати обсяг початкового опису.

Одними із високоінформативних стислих подань для багатовимірних даних є проекції [33]. Функція проекції для бінарних даних обчислюється як сума двійкових компонентів за фіксованим напрямком (фактично це кількість одиниць або нулів).

Застосуємо просторове оброблення і на підставі одно-бітових проєкцій побудуємо інтегровану в межах окремих фрагментів систему ознак. Для матриці D опису обчислимо значення проєкції вздовж стовпців (1-бітовий фрагмент), отримаємо вектор $t = (t_1, \dots, t_j, \dots, t_n)$, де $t_j = \sum_{i=1}^s d_{i,j}$, $j = \overline{1, n}$ (фактично t_j – це сума одиниць у стовпці). Це і будуть ознаки найнижчого рівня, що у повній мірі зберігають інваріантні властивості множини дескрипторів Z .

Тепер на базі проєкційних ознак t_j обчислимо ознаки більш високого рівня у вигляді атрибутів u_k для блоків у вигляді сукупностей стовпців [14, 20, 88]

$$u_k = \sum_{j=k}^{k+b-1} t_j, \quad (3.19)$$

де $b = n / m$ – розмір, а $k = 1, b + 1, 2b + 1, \dots, n - b + 1$ – номер фрагмента.

Ознаки (3.19) є варіантом крос-кореляційного оброблення з прямокутною маскою розміром $b \times s$, що містить одиниці. У результаті обчислення (3.19) отримуємо цілочисельний вектор u_k розмірністю m . Параметр m є характеристикою системи фрагментів, він зменшується від n до 1 зі збільшенням розміру фрагмента від 1 до n .

Значення вектора $u = (u_1, \dots, u_k, \dots, u_m)$ можуть бути використані як самостійні структурні ознаки статистичного виду, вони обчислюються на підставі побітового аналізу стовпців матриці D . Опрацювання (3.19) залежить від параметра b та реалізує просторове оброблення (інтегрування даних) на множині дескрипторів. Зважаючи на таку просту модель обчислення функцій проєкцій, всі вони для довільного розміру фрагмента визначаються нескладно (логічно або додаванням цілих чисел), базуючись на значеннях 1-бітових проєкцій, що є статистичним розподілом даних.

На основі подання (3.19) може бути застосовано також ієрархічний спосіб розпізнавання, що використовує для зіставлення з еталонами систему ознак u_k з різним ступенем інтегрування даних, що визначається розміром блока.

Діапазон цілих значень для ознак u_k безпосередньо можна визначити за розміром фрагмента як $u_k \in \{0, \dots, sb\}$. Загалом вектор u є інтегрованою характеристикою опису, його можна окремо нормувати за кількістю дескрипторів або розміром фрагменту.

Обсяг значень для ознак (3.19) змінюється від одного числа до розміру n дескриптора і зменшується зі зменшенням розміру фрагмента. При $m=1$ маємо $b=n$, тобто кожний дескриптор містить один фрагмент, а величина u_1 визначається одним значенням і дорівнює сумі одиничних бітів усього опису. При $m=n$ маємо іншу граничну ситуацію – n 1-бітових фрагментів, а u_k обчислюється як сума одиничних бітів кожного стовпця матриці D .

Модель оброблення (3.19) реалізує процедуру зменшення інформаційної надмірності просторового сигналу за рахунок допустимого зниження (з точки зору якості розрізнення) ступеня розподільної здатності ознакового подання опису.

Оброблення виду (3.19) є різновидом просторового аналізу значень для множини дескрипторів вздовж координат їх компонентів. Іншими варіантами такого оброблення є фільтрація чи розклад вектору дескриптора за системою ортогональних функцій.

Зауважимо, що кількість одиниць у фрагменті бінарного дескриптора – теж ознака, що може бути використана для розпізнавання. Якщо додатково виконати ранжирування фрагментів за кількістю одиниць чи нулів, то отримаємо кортеж, інваріантний до їх порядку слідування.

Одним із шляхів спрощення аналізу блокових даних є використання значень окремих параметрів розподілів Q . Визначення спільного дескриптора (СД) [4], який можна ще назвати статистичним центром, для

побудови концентрованого опису зображення можна базувати, наприклад, на застосуванні найбільш вживаних шаблонів – бінарних масок (а фактично – цілих чисел), значення яких найчастіше зустрічаються у системі фрагментів.

Визначимо максимальне значення розподілу фрагмента у кожному стовпці матриці Q і сформуємо ланцюжок СД d як результат зчеплення найбільш вживаних значень у межах опису

$$d_v = \arg \max_{i=1, \dots, w} q_{i,v}, v = \overline{1, m}, \quad (3.20)$$

$$d = d_1 \& d_2 \& \dots \& d_m,$$

а значення d_v як аргумент максимуму отримано для кожного фрагмента. Таким чином, СД для опису Z тут має вигляд вектору, що містить значення, які найчастіше зустрічаються у відповідному місці кожного із дескрипторів.

У спрощеній моделі (3.20) варто враховувати лише достатньо вагомі за максимумом величини із граничною характеристикою значимості $q^* = \max_{i=1, \dots, w} q_{i,v}$ наприклад, за умови $q^* \geq 0,5$ або $q^* \geq 0,25$. Можна на попередньому етапі проаналізувати і сформувати список фрагментів, які для конкретного еталону мають високий рівень підтримки q^* всередині опису. Решту фрагментів доцільно виключити із аналізу.

Обмеження на значення підтримки забезпечує адаптацію створеного СД до аналізованого масиву даних у плані точнішого відтворення його вмісту. Продукція виду (3.20) відповідає принципу максимуму апостеріорної ймовірності для системи блоків. Для суттєво різних описів такий аналіз приносить необхідний результат [4].

Проаналізуємо способи обчислення релевантності описів на основі просторового оброблення опису. Один із загальноприйнятих способів визначення релевантності описів у запропонованому просторі статистичних ознак полягає в обчисленні відстані між матрицями Q розподілів або

побудованими векторами ознак (3.19). Шляхом оптимізації значення відстані на множині еталонів визначається клас розпізнаваного об'єкта. Цей спосіб безпосередньо використовує інтегровані дані про розподіли еталонів та розпізнаваного об'єкта.

Інший спосіб, який застосовують при низьких відношеннях сигнал-шум, полягає у конкурентному віднесенні кожного із дескрипторів аналізованого опису візуального об'єкта до одного із узагальнених подань еталонних описів, який використовується замість матриці еталонних даних [28]. Потім підсумовується кількість дескрипторів КТ об'єкта, віднесених до кожного із еталонів, і визначається еталон з найбільшим числом голосів.

Вектор ознак $u = (u_1, \dots, u_k, \dots, u_m)$ безпосередньо можна використати у першому способі на підставі обчислення манхеттенської відстані у векторному просторі цілих чисел

$$\rho[u(o), u(e)] = \sum_{l=1}^m |u_l(o) - u_l(e)|, \quad (3.21)$$

де $u(o), u(e)$ – ознаки розпізнаваного об'єкта та еталону.

Значення міри (3.21), очевидно, зростає поетапно по мірі досягнення кінцевої суми. На підставі цього факту можна впровадити способи подальшого скорочення обсягу обчислень, зупинивши процес при досягненні сумою в (3.16) деякого критичного порогу щодо відмінності описів [26, 30].

На основі ансамблю мір (3.21) при різноманітних значеннях $m = 1, \dots, n$ можна побудувати ієрархічну міру, що враховує весь діапазон розподільчої здатності запропонованої системи ознак.

Для другого способу визначення релевантності треба спочатку нормувати еталонну систему ознак $u(e)$, зважаючи на параметри числа дескрипторів та фрагментів, для яких вона отримана. Далі для кожного дескриптора аналізованого об'єкта необхідно сформулювати відповідно нормований до вибраного діапазону вектор $u(o)$. Для конкурентного

віднесення дескриптора до одного з еталонів також можна застосувати відстань (3.21) [9, 13, 14].

Зважаючи на властивість модуля чисел $|a + b| \leq |a| + |b|$ та інтегрований спосіб побудови ознак u , можна сказати, що значення метрики (3.21) не зростає зі зменшенням m . Це свідчить про те, що зі зменшенням m відмінність між різними зображеннями у такому просторі ознак не збільшується, а відповідно ступінь їх розрізнення може погіршуватися. Зниження розподільної здатності системи розпізнавання тут пояснюється зростанням ступеня узагальнення чи інтегрованості даних u .

Альтернативою для критерію (3.16) щодо визначення ступеня відмінності інтегрованих описів u є застосування метрики для порівняння множин [28]. Прикладом є метрика Танімото, яку можна впровадити для зіставлення відповідних фрагментів, а в подальшому інтегрувати для усього набору фрагментів. Метрика Танімото для множин A , B визначається співвідношенням $t[A, B] = \frac{A \Delta B}{A \cup B}$, де чисельник містить симетричну різницю, а знаменник – об'єднання множин. Якщо позначити $c(A) = \text{card}(A)$ – число елементів скінченної множини, то для метрики $t[A, B]$ отримуємо

$$t[A, B] = \frac{c(A \setminus B) + c(B \setminus A)}{c(A) + c(B) - c(A \cap B)}, \quad (3.22)$$

де в чисельнику стоїть число елементів для різниць множин.

Гуртову за системою фрагментів відстань Танімото обчислюємо як

$$t[u(o), u(e)] = \sum_{k=1}^m t[u_k(o), u_k(e)], \quad (3.23)$$

де доданки визначаються за моделлю (3.14).

Значення виразу (3.23) має властивості метрики як сума метрик [29].

Приклад. Нехай маємо подання бінарного ланцюжка $a(a_0, a_1)$ як проекції у вигляді кількості нулів a_0 та кількості одиниць a_1 . Для двох різних бінарних ланцюжків $a=001100$, $b=101010$ маємо описи проекцій: $a(4,2)$ та $b(3,3)$. Нормована відстань (3.21) між проекціями ланцюжків складає $2/12$ (у знаменнику – максимальне значення), а відстань (3.23) дорівнює $2/7$. Зважаючи на те, що ці ланцюжки за відстанню Хемінга різняться у трьох бітах, відстань (3.21) краще відображає відмінності у їх властивостях, бо більша за значенням.

Просторовий аналіз даних як множини фрагментів отримав поширення у теорії комп'ютерного зору [26, 30]. Він забезпечує можливість прийняття рішень за окремими елементами візуальних об'єктів, що важливо у прикладних застосуваннях. При цьому управління такими характеристиками блочної системи, як значення порогів на число, подібність чи значимість фрагментів, покладається на користувача. Встановлення оптимального розміру фрагмента, порогу для визначення еквівалентності фрагментів, порогу для встановлення значущості чи інформативності опису фрагмента треба виконувати, виходячи із прикладної бази зображень, де здійснюється розпізнавання.

Подання (3.19) розкриває нові можливості для побудови моделей класифікації на підставі системи фрагментів. Кожний з m різних «маленьких» блоків-еталонів може незалежно голосувати за відповідний клас. Апаратно цей процес можна реалізувати паралельним способом. Далі вирішальна функція голосів $G(g_1, g_2, \dots, g_m)$ окремих фрагментів визначає найбільш ймовірний клас, до якого й відносять аналізований дескриптор чи СД об'єкта. Таким чином, застосовується структура всередині системи дескрипторів, де окремі фрагменти підтримують або ні класифікаційне рішення.

На основі моделі (3.19) може бути застосовано також ієрархічний спосіб розпізнавання, що використовує для зіставлення з еталонами систему

ознак u_k з різним ступенем інтегрування даних, що визначається розміром блока [20].

Діапазон цілих значень для ознак u_k безпосередньо можна визначити за розміром фрагмента як $u_k \in \{0, \dots, sb\}$. Загалом вектор u є інтегрованою характеристикою опису, його можна окремо нормувати за кількістю дескрипторів або розміром фрагменту.

Обсяг значень для ознак (3.19) змінюється від одного числа до вектора розміру n (розмір дескриптора) і зменшується зі збільшенням розміру фрагмента. При $m=1$ маємо $b=n$, тобто кожний дескриптор містить один фрагмент, а величина u_1 визначається одним значенням і дорівнює сумі бітів усього опису. При $m=n$ маємо іншу граничну ситуацію: n 1-бітових фрагментів, а u_k обчислюється як сума стовпця матриці D .

Модель (3.19) реалізує процедуру зменшення інформаційної надмірності просторового сигналу за рахунок допустимого зниження (з точки зору якості розрізнення) ступеня розподільної здатності системи ознак подання опису.

Оброблення виду (3.19) є різновидом просторового аналізу значень для множини дескрипторів вздовж координат їх компонентів. Іншими варіантами такого оброблення є фільтрація чи розклад дескриптора за системою ортогональних функцій.

Релевантність r описів a та b визначимо на підставі зіставлення (3.21) розподілів на підставі манхеттенської відстані між векторами $u(a), u(b)$. Для більш точного оцінювання значення релевантності можна застосувати нормовані значення

$$r^* = \frac{r}{r_1}, \quad (3.24)$$

де r_1 – максимально можлива відстань для обраного числа дескрипторів і ієрархічної системи фрагментів.

Висновки

1. Система статистичних розподілів структурного опису дає можливість побудувати принципово новий простір ознак зображень. Фрагменти дескрипторів ключових точок можна розглядати як систему просторових ознак, застосування якої спрощує оброблення та зберігає чутливість щодо розрізнення даних.

2. Впровадження блокової структури даних дає можливість гнучко керувати процесом визначення релевантності описів. Бітовий вид дескрипторів приваблює можливістю застосування засобів нескладного за обчислювальними витратами бінарного оброблення з відповідним запровадженням у прикладних задачах.

3. Застосування різноманіття статистичних критеріїв дало ідентичні обґрунтовані висновки про значущість відмінностей описів візуальних об'єктів, що підкреслює об'єктивність проведеного дослідження.

4. Просторові методи оброблення сигналів дають можливість шляхом опрацювання чи аналізу підструктур даних виявити статистичні закономірності та стійкі при розпізнаванні характеристики об'єктів.

4. Результати комп'ютерного моделювання розроблених методів

4.1. Експериментальний аналіз зіставлення розподілів фрагментів

Розроблений метод обчислення релевантності застосовано на прикладі зображень ікон («Казанська богоматір», «Секстинська мадонна» і т.п.) розміром 400x540 у середовищі Visual Studio 2017 із використанням засобів бібліотеки Open CV [86], ілюстрацію зображення показано на рис. 4.1. Застосовано дескриптори BRISK розмірністю $n=512$.



Рис. 4.1. Зображення ікони

На рис. 4.2 наведено приклад сформованого розподілу щодо фрагментів дескрипторів опису для двох ікон (синій та жовтий колір) за першими двома бітами.

Як бачимо із рис. 4.2, розподіли досить схожі, що можна пояснити однотипним характером зображень. Для інших, вельми різних зображень (наприклад, ікона і герб м. Харкова), в експерименті розподіли за двома бітами більш суттєво різняться.

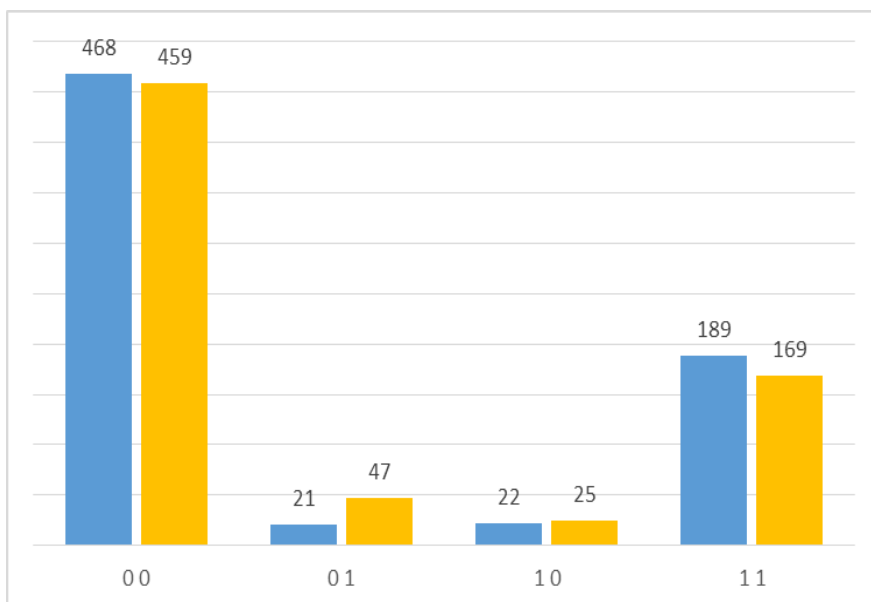


Рис. 4.2. Приклад розподілів за значеннями двох бітів

Експериментальна табл. 4.1 містить величини нормованої манхеттенської відстані r^* (3.21) для описів зображень двох ікон в залежності від числа дескрипторів (16...700) у складі опису та кількості бітів (1...8) у фрагментах розподілу. Міра подібності за голосуванням (підрахунок подібних дескрипторів з порогом еквівалентності 20% від максимуму значення відстані Хемінга) для цих двох ікон склала 0,12. Зважаючи на це, із табл. 4.1 бачимо, що запропонована міра (3.21) має достатні можливості для розрізнення таких близьких за змістом зображень, як ікони. Зі зростанням числа бітів розподілу від 1 до 8 відстань збільшується, тобто властивість розпізнавання покращується.

У той же час зі зменшенням числа КТ опису відстань також збільшується, але при цьому можна висловитись про зменшення достовірності розпізнавання, так як два зовсім різних зображення можуть мати на невеликих вибірках ознак як суттєву подібність, так і значну відмінність. Значення табл. 4.1 показують, що 50-100 КТ достатньо для розрізнення зображень ікон. Для інших типів зображень або кількості розпізнаваних еталонів ці значення можуть різнитися.

Таблиця 4.1. Значення відстані r^* для різної кількості дескрипторів та бітів

Кількість дескрипторів	Число бітів розподілу			
	1	2	4	8
700	0,059	0,084	0,143	0,319
500	0,061	0,086	0,149	0,326
100	0,071	0,110	0,204	0,450
50	0,106	0,165	0,290	0,570
30	0,108	0,175	0,320	0,600
16	0,127	0,210	0,402	0,699

За результатом порівняння значень табл. 4.1 з відповідними значеннями, отриманими для суттєво різних за змістом зображень (ікона і герб), виявилось, що значення у стовпці для 1 біта приблизно на 50% вище, а у стовпці для 8 бітів – на 20% більші. Цей факт підтверджує значимі властивості застосованої моделі даних розподілів щодо розрізнення об'єктів навіть у випадку незначної кількості бітів фрагмента.

Проаналізуємо наявні обчислювальні витрати для реалізації розглянутих методів у порівнянні з традиційними підходами, час обчислення з використанням комп'ютерних моделей у секундах зафіксовано у табл. 4.2 [12, 17].

Таблиця 4.2. Оцінки часу при комп'ютерному моделюванні мір

Кількість дескрипторів	Міра подібності (3)	Число бітів розподілу			
		1	2	4	8
700	10,7	0,008	0,011	0,028	11
500	6,4	0,006	0,01	0,017	6,2
100	0,48	0,002	0,002	0,007	1
50	0,11	0,001	0,002	0,004	0,915
30	0,08	0,001	0,001	0,003	0,527
16	0,045	< 0,001	< 0,001	0,002	0,315

Як можна побачити із табл. 4.2, запропонований підхід на основі розподілів фрагментів має суттєві часові переваги порівняно з традиційним підходом, так як порівняльний час для визначення релевантності для них

зменшується приблизно у 1300 разів для 1-бітового подання та у 400 разів для 4-бітового. При цьому час для 8-бітового варіанту майже дорівнює затратам для традиційного методу, однак, основні витрати при цьому ідуть суто на формування розподілів, а не на обчислення релевантності. Перспективними тут є застосування моделей для визначення релевантності без побудови розподілів для розпізнаваного зображення, спираючись виключно на еталонні розподіли, побудовані на етапі попереднього оброблення [16]. Зрозуміло при цьому, що часові витрати для усіх моделей обчислення релевантності залежать від режиму виконання програми та пропорціональні числу дескрипторів опису.

4.2. Розрахункова оцінка ступеня відмінності експериментальних даних

Наведемо наші міркування на користь твердження щодо підвищення рівня розрізнення описів об'єктів при подовженні ланцюжків у розподілах, виходячи з порівняльного аналізу значень цих розподілів відповідно до окремих фіксованих ланок. Для розрахунків на підставі даних наведеного експериментального прикладу у випадку $s = 700$ дескрипторів нами здійснено перехід до нормованих значень розподілів шляхом ділення на кількість значень описів за фрагментами окремих ланок. Наприклад, для опису першої ікони при розподілі пар бітів комбінація 00 (ланка розподілу) загалом зустрічалася 66507 разів у множині із 256 пар, а комбінація 01 – 24414 разів. Ці дані і застосовувалися для нормування розподілів. У результаті сума нормованих значень для кожної з комбінацій дорівнювала 1.

Для конкретності викладок назвемо X – значення розподілу для першої ікони, Y – для другої. Аналіз результатів обчислень за різними статистичними критеріями показав наступне.

Значення парного коефіцієнта кореляції між показниками X та Y поступово знижується із зростанням кількості бітів у фрагменті. При цьому, наприклад, для 30% типів ланок з числа усіх можливих у випадку 8-ми бітового розподілу маємо значення корельованого зв'язку, нижче за помірне (згідно шкали Чеддока, тобто коефіцієнт кореляції не перевищує 0,5). Зокрема, результат зіставлення кількості ланок типу 10101000 за відповідними фрагментами порівнювальних об'єктів демонструє досить слабкий корельований зв'язок з коефіцієнтом кореляції 0,25. Цей результат принципово відрізняється від аналогічних розрахунків для одно- і дво- бітового розподілів, які свідчать про високий або дуже високий кореляційний зв'язок (коефіцієнт кореляції приймає значення від 0,74 і вище).

Значення інформаційної ентропії Шеннона [85] окремо по кожному показнику X та Y очікувано зменшується при збільшенні довжини фрагмента, що означає відповідне підвищення інформативності досліджуваних даних. У табл. 4.3 наведений результат таких розрахунків для ланок 0, 00, 0000.

Таблиця 4.3. Значення інформаційної ентропії Шеннона

Об'єкт	Кількість бітів розподілу		
	1	2	4
Ікона 1	6,181	5,443	4,705
Ікона 2	6,151	5,403	4,656

Цікавим представляється також дослідити значення міри подібності між показниками X та Y, в якості якої можна застосувати α -розходження Рен'ї при $\alpha=0,5$, яке в певному сенсі є оптимальним серед припустимих значень $\alpha \in [0;1]$ [9, 73]. Значення α -розходження Рен'ї при $\alpha=0,5$ для аналогічних табл. 3 ланок ікони 1 складають 0,0056; 0,0084; 0,0109.

Ці розрахунки демонструють передбачуване збільшення значення міри подібності із зростанням довжини фрагментів розподілу, що знову

підтверджує факт покращення рівня розрізненості досліджуваних об'єктів із подовженням ланцюжків у розподілах.

4.3. Результати моделювання засобів просторового оброблення

Для оцінювання ефективності застосування запропонованих підходів просторово-статистичного оброблення даних структурних описів зображень розроблено програмну модель у середовищі Visual Studio 2017 із використанням засобів бібліотеки Open CV [86]. Дослідження проведено для зображень ікон і гербів міст України розміром 400x540. Застосовано детектор BRISK з розмірністю дескриптора $n=512$. Змінюване число дескрипторів у аналізованому описі у 700, 500, 100, 50 КТ формувалось шляхом випадкового вибору із початкової їх множини, отриманої детектором.

Рис. 4.3, 4.4 порівняльним чином містять значення ознак u_1, \dots, u_5 зображень ікони (жовтий колір) та герба Харкова (синій колір) для 1-бітового (рис. 4.3) та 8-ми бітового подань (рис. 4.4). Таблиця 4.4 включає значення манхеттенської відстані для повних описів ікони та герба у межах запропонованої системи ознак u , нормованої на загальне число sb аналізованих бітів.

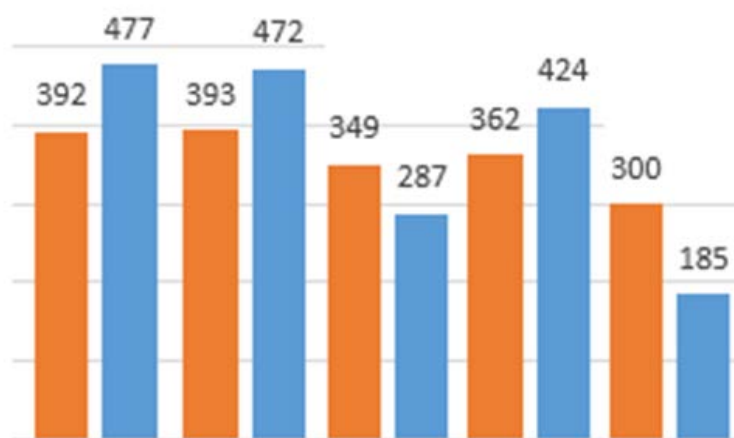


Рисунок 4.3. Порівняльні значення ознак u_1, \dots, u_5 для 1-бітового подання

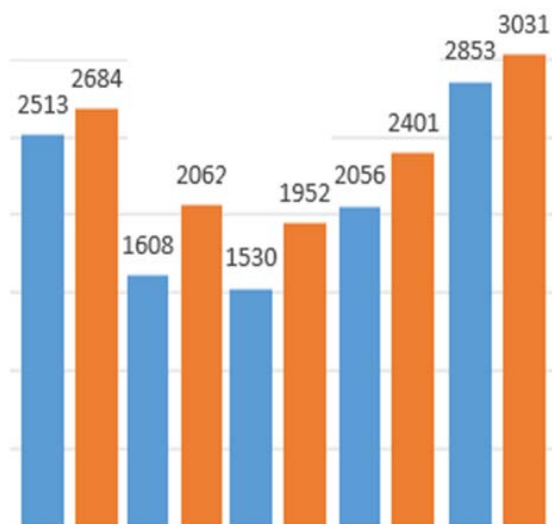


Рисунок 4.4. Порівняльні значення ознак u_1, \dots, u_5 для 8-бітового подання

Таблиця 4.4. Значення манхеттенської відстані для різнотипних зображень

Число дескрипторів опису	Загальне число бітів для ознаки u			
	1	2	4	8
700	0,096	0,085	0,071	0,063
500	0,070	0,061	0,051	0,047
100	0,015	0,013	0,011	0,009
50	0,008	0,007	0,006	0,005

За результатами моделювання зазначимо, що для однотипних зображень (дві різні ікони) обчислена відстань приблизно менша на 50%, що підкреслює достатню чуттєвість запропонованих ознак до вмісту зображення. Незважаючи на отримані у таблиці 4.4 незначні абсолютні значення відстані, експерименти показали безпомилковий результат класифікації як окремо у межах бази із 10-ти ікон та гербів міст України, так і у змішаній базі із 20 зображень.

Час обчислення відстані для аналізованих варіантів оброблення практично не відрізняється і навіть для кількості у 700 точок знаходиться у межах 1 мілісекунди, що наголошує обчислювальну простоту і достатньо високу швидкодію запропонованих методів оброблення. Більш детальний аналіз наявних обчислювальних витрат для визначення відстані при $s = 700$ показав, що для 8-ми бітових ознак час приблизно у 5 разів менший, ніж для однобітових, що говорить про можливість подальшого покращення характеристик швидкодії. Зважаючи на те, що у порівнянні з традиційним методом обчислення релевантності з використанням моделі голосування дескрипторів час обчислення релевантності у просторі проєкційних ознак загалом скорочується у тисячі разів [16], можна зробити висновки про перспективність застосування запропонованих ознакових моделей.

Як бачимо за даними таблиці 4.4, значення міри релевантності описів у створеному просторі ознак прогнозовано зменшується зі зменшенням числа точок, що вказує на зниження розподільної здатності системи розпізнавання і можливого зниження рівня безпомилкового розпізнавання. Наші експерименти показали, що для таких інформаційно насичених зображень, як ікони та герби, достатнім для безпомилкової класифікації є число точок, що перевищує 100.

Із даних таблиці 4.4 також бачимо, що зі зростанням ступеня інтегрованості, що відповідає збільшенню числа бітів від 1 до 8, розподільчі властивості ознак u дещо знижуються, але незначним чином. Так, відстань для 8-бітових ознак при $s = 700$ склала величину 0,063 у порівнянні з 0,096 для 1-бітового подання. Із отриманих експериментальних результатів можна зробити висновок про можливість практичного використання ознак u із зазначеного діапазону числа бітів: 1, 2, 4, 8.

У порівнянні з ознаками розподілів у необробленому вигляді запропонована система просторових ознак u має суттєво простіше обчислення та практично ті ж показники результативності класифікації, тому

може бути успішно застосована у прикладних задачах задля прискорення розпізнавання.

4.4 Аналіз комп'ютерного моделювання ієрархічних ознак

Розроблену модель побудови ієрархічних ознак застосовано на прикладі зображень гербів університетів Оксфорду, Гарварду, Пекіну та Массачусетса. Розмір зображень складає 600x600. Програмне середовище – Visual Studio 2015 із використанням засобів бібліотеки Open CV [86]. Ілюстрація зображення герба Гарварду та Массачусетса показана на рис. 4.5. Для моделювання застосовано дескриптори ORB розмірністю $n = 256$.

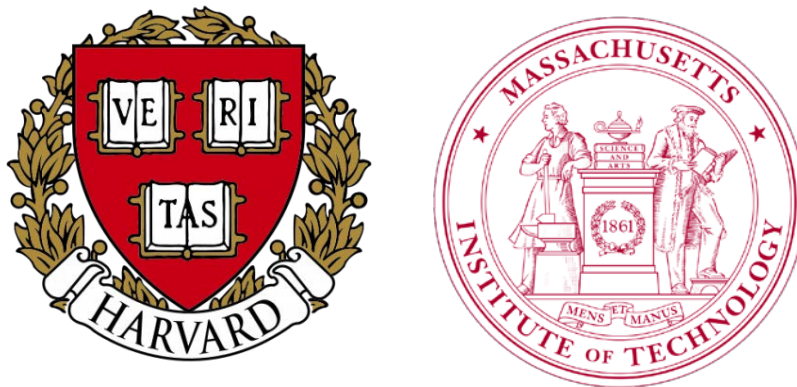


Рис. 4.5. Зображення гербів

На рис. 4.6 наведено приклад сформованого 1-бітового розподілу (256 ланок дескриптора ORB) для числа нулів у 100 дескрипторах герба рис 4.5.

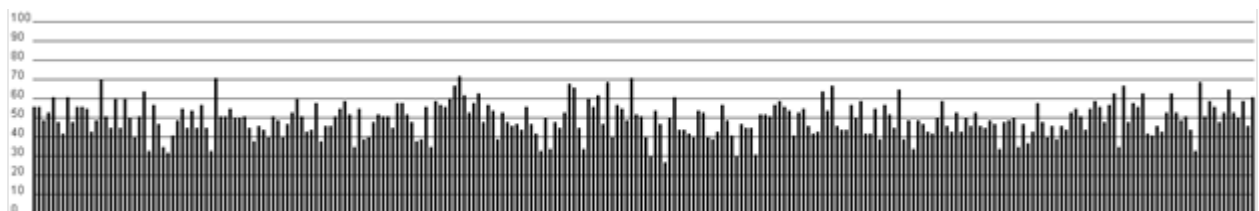


Рис. 4.6. Приклад розподілу для 100 дескрипторів У таблиці 4.5 наведено значення манхеттенської відстані r для 1-бітових розподілів щодо 100 дескрипторів кожного із чотирьох гербів.

Таблиця 4.5. Відстані між 1-бітовими розподілами для 100 дескрипторів

	Оксфорд	Гарвард	Масачуссет	Пекін
Оксфорд	0	1922	3782	5232
Гарвард	1922	0	2868	4182
Масачуссет	3782	2868	0	2962
Пекін	5232	4182	2962	0

На рис. 4.7 показано схему побудови ієрархічних ознак для емпіричного опису зображення рис. 4.5 у складі 100 дескрипторів, де значення у нижніх колах відповідають 8-бітовому розподілу нулів, а верхнє значення дорівнює загальному числу нулів опису. Аналогічну структуру можна побудувати для довільного опису.

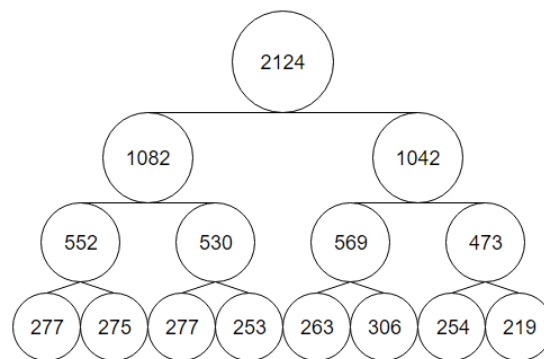


Рис. 4.7. Значення ієрархічних ознак на прикладі герба Гарварду (рис. 4.5)

Таблиця 4.6 містить значення відстані між розподілами зображень гербів Масачусетського та Оксфордського університетів для 100, 300, 500 дескрипторів та різним числом бітів у фрагментах розподілу. Також у таблиці 4.6 наведені відстані для скорочених розподілів цих гербів, де вибрано кожен другий та четвертий біти із однобітового розподілу.

При зменшенні кількості ключових точок майже пропорційно зменшуються й максимальна та фактична відстань між розподілами. Тобто збільшення числа дескрипторів від 100 до 300 не має суттєвого впливу на характеристики розпізнавання зображень. Значення таблиці 4.6 показують, що 100 ключових точок достатньо для розрізнення зображень гербів. Для

інших типів зображень або кількості розпізнаваних еталонів ці значення можуть відрізнятися.

У таблиці 4.7 наведено дані наявних обчислювальних витрат для реалізації варіантів розглянутих моделей при обчисленні релевантності описів зображень гербів Масачуссета та Оксфорду в залежності від кількості бітів у фрагментах розподілу. Таблиця 4.7 містить час опрацювання комп'ютерними моделями у наносекундах.

Аналізуючи значення таблиці 4.6 та оцінку часу оброблення з таблиці 4.7 для 100 КТ, можна сказати, що при обчисленні по правилу «кожен другий» час прогнозовано зменшився у порівнянні з 1-бітовим розподілом приблизно у 1,5 рази. А нормована відстань дещо зменшилась від 0,147 до 0,105.

Таблиця 4.6. Значення нормованої відстані r^* для різної кількості дескрипторів та бітів розподілу

Число дескрипторів	Число бітів розподілу							
	1	2	4	8	128	256	Кожен 2-й	Кожен 4-й
500	0,142	0,098	0,073	0,060	0,046	0,046	0,138	0,144
300	0,139	0,095	0,071	0,061	0,045	0,045	0,147	0,147
100	0,147	0,102	0,074	0,061	0,046	0,046	0,105	0,155

Таблиця 4.7. Оцінка часу при обчисленні релевантності

Число дескрипторів	Число бітів розподілу							
	1	2	4	8	128	256	Кожен 2-й	Кожен 4-й
500	66929	33944	26222	22095	20388	19913	32493	26511
300	46465	65172	34177	22327	19612	19914	33139	26087
100	47018	32586	24540	21472	18572	16293	32133	25384

При обчисленні по правилу «кожен четвертий» час зменшився у порівнянні з 1-бітовим розподілом у 1,9 рази. А нормована відстань r^* практично не змінилась в межах 0,15. Ці показники стверджують про

достатні властивості модифікацій побудованих ознак для розпізнавання зображень гербів зі значно меншим часом оброблення.

Відмітимо, що оцінка часу у таблиці 4.7 враховує тільки знаходження відстані, не включаючи затрати на обчислення розподілу, тобто програма орієнтована на порівняння з еталонними розподілами, побудованими на етапі попереднього оброблення. У цьому випадку, як видно із таблиці 4.7, не має суттєвого значення кількість КТ, а важлива тільки кількість бітів у фрагментах розподілу. Чим більше використано бітів – тим повільніше працює програма. Як можна побачити із таблиці 4.7, з кожним збільшенням числа бітів у розподілі у 2 рази час виконання зменшується приблизно у 1,3 рази. Зрозуміло, що, якщо брати до уваги час, затрачений на обчислення розподілу, то він прямо пропорційний числу дескрипторів опису.

Зважаючи на те, що час обчислення релевантності за розподілами приблизно у 1000 разів менший, ніж для традиційних процедур голосування, бачимо, що за допомогою впровадження системи ієрархічних ознак можна добитися подальшого скорочення часу обчислень, забезпечуючи при цьому достатньо високу результативність розпізнавання.

Ступінь інтегрованості призводить до зменшення рівня відмінності ознак від 0,147 (для 1-бітового) до 0,06 для 8-бітового, що зменшує розрізненість описів у інтегрованому поданні. Але і для 256 бітів спостерігаються відмінності між ієрархічними описами гербів, що дозволяє їх розпізнавати між собою. При цьому час для 256 бітів зменшився майже в 3 рази.

Висновки

1. Узагальнюючий статистичний аналіз значень фрагментів для множини дескрипторів, що складають структурний опис зображення, є перспективним напрямком скорочення обчислювальних витрат на розпізнавання. Ймовірніше подання даних сприяє не тільки суттєвому прискоренню процедур оброблення (більш ніж в тисячу разів для

розглянутих прикладів), але й забезпечує достатній рівень розрізненості образів.

2. Експерименти показали, що впроваджена блочна система створює достатньо репрезентативний простір, так як семантично пов'язані зображення у ньому залишаються досить подібними між собою.

3. Трансформація ймовірнісних розподілів для фрагментів дескрипторів і зіставлення образів у просторі ієрархічних ознак забезпечують необхідну результативність розпізнавання. Таке оброблення реалізується значно швидше, ніж у просторі розподілів та у процедурах голосування. Оброблення за значеннями згрупованих розподілів формує ієрархічну структуру даних зі змінюваним параметром деталізації та значимими властивостями опису.

4. Практичні рекомендації із проведеного дослідження полягають у результативному застосуванні інтегрованих ознак для груп бітів як ефективного засобу з точки зору швидкодії обчислень.

5. Результати моделювання підтверджують суттєвий рівень підвищення швидкодії при обчисленні релевантності, результативність запропонованих ієрархічних ознак на прикладах зображень. Розроблено прикладні програмні моделі для дослідження та впровадження методів класифікації в системах комп'ютерного зору.

5. Прикладні системи статистичних ознак для структурних описів

Цей розділ містить матеріали досліджень із використання статистичних підходів для обчислення практичних систем ознак на підставі використання детектора SURF, що формує опис у вигляді множини дескрипторів – числових векторів із простору R^n . Опис $Z \subset R^n$ представляють скінченною мультимножиною дескрипторів. Детектор SURF формує опис як підмножину $R_1^n = \{z \mid z \in R^n, \|z\| \approx 1\}$ n -вимірних дійсних векторів з наближеною до одиниці евклідовою нормою: $\|z\| = \sqrt{\sum_{i=1}^n z_i^2} \approx 1$ [24, 25].

Так як число класів візуальних об'єктів у базі даних може бути значним (тисячі еталонів або відео кадрів), то час, необхідний для пошуку відповідності нового образу шляхом порівняння з кожним із еталонів бази може виявитися значним. Дієвим способом підвищення швидкодії оброблення при порівнянні описів у вигляді множин є перехід до зіставлення векторів, що характеризують ці множини. Для цього потрібно, наприклад, шляхом кластеризації множини дескрипторів КТ сформувати векторний опис [11, 21].

Ефективним способом, застосовуваним у математичній лінгвістиці для пошуку документів за їх змістом, є вживання показників частоти появи слів із документа пошуку і словника, на основі якого побудовано зміст бази слів [50, 51]. Його відносять до технології «мішок слів» [24, 49]. До таких методів відносять tf-idf, який представлений рядом модифікацій [51]. Розпізнавання за множиною структурних ознак подібно лінгвістичному підходу не вимагає врахування послідовності їх формування. Може виявитися плідною ідея побудувати розпізнавання так, щоб дескриптори зображення уявити як «візуальні слова» в сформованому словнику, а лінгвістичні характеристики об'єкта – у вигляді вектора. Тоді зображення еталонів і об'єкту буде представлено векторами, що значимо знижує витрати на розпізнавання [24].

Аналіз лінгвістичних властивостей дає також можливість оцінити, наскільки часто зустрічається конкретний дескриптор в структурному описі, і зосередити пошук на рідкісних і, отже, найбільш інформативних ознаках. Попередньо із опису можна відфільтрувати ознаки, що досить часто зустрічаються, вони не завжди корисні для прийняття рішень.

Поряд із цим інформативними можна вважати значення відстаней між дескрипторами, що складають опис, тоді як повноцінною інтегральною характеристикою може бути деяка функція від сукупності відстаней [24]. З цієї точки зору звернемо увагу на таку стрижневу характеристику, як медіана множини [53]. Медіанна обробка досить результативна в плані скорочення обсягу обчислень, так як дозволяє трансформувати опис до списку елементів із закріпленими номерами у вигляді вектора дескрипторів [25].

Поставимо мету вивчити можливості та оцінити ефективність застосування в структурному розпізнаванні лінгвістичної схеми зважування *tf-idf* і принципів медіанної обробки на множині дескрипторів КТ в плані побудови векторного уявлення опису. Обчислення лінгвістичних і медіанних інтегральних характеристик дозволяє реалізувати стиснення до векторного вигляду уявлення еталонів і об'єкта, що загалом значно покращує швидкодію розпізнавання.

Завданнями дослідження є вивчення особливостей побудови вектора-ознак *tf-idf* і списку дескрипторів пост-медіанної обробки стосовно множин структурних ознак зображень, а також оцінювання ефективності вживання векторного порівняння для прикладних баз зображень.

5.1. Побудова схеми зважування *tf-idf*

Метод *tf-idf* застосовують у задачах інформаційного пошуку при оцінюванні змісту текстових документів, складених з фіксованого числа термінів – словника [51]. Відповідно до методу *tf-idf* обчислюють числовий

вектор, компонентами якого є ваги термінів. У результаті кожен документ отримує опис у вигляді вектора розміром встановленого обсягу словника. Це дозволяє порівнювати між собою різноманітні за змістом і розміром документи і отримувати для них значення релевантності (подібності або відповідності).

Спочатку необхідно сформувати словник – скінченну базу термінів, в рамках якої реалізується лінгвістичний аналіз. Потім для кожного терміна словника обчислюється частота терміна в документі (tf – term frequency) і зворотна документна частота (idf – inverse document frequency). Модель обробки у методі tf-idf називають мішком слів (bag of words model). В її рамках порядок проходження термінів у документі ігнорується, а основна увага приділяється кількості входжень кожного терміна. Інтуїтивно ясно, що два документи з однаковими «мішками слів» за змістом подібні. Використання зворотної документної частоти idf необхідно для ослаблення впливу на оцінку релевантності термінів, що часто зустрічаються. Вага tf-idf терміна визначається як

$$g_i = \frac{a_i}{c} \cdot \log_2 \left(\frac{b}{d_i} \right), \quad (5.1)$$

де a_i – число зустрічальності i -го терміна у документі об'ємом із c термінів, b – кількість документів у колекції, d_i – число документів із їх загального числа b , в яких зустрічається i -тий термін, \log_2 – логарифм числа за основою два [51]. Функція tf-idf реалізує інтуїтивне уявлення, що чим частіше терм зустрічається в документі, тим точніше він описує його зміст, і чим в більшій кількості документів зустрічається терм, тим менше розпізнавальних властивостей він має.

У літературі [50] досліджено велике розмаїття варіантів методу tf-idf для прикладних задач. Основні ідеї модифікацій зводяться до побудови

евристичних правил для ранжирування та порогового оброблення вектора $tf\text{-}idf$ з метою управління вагами термінів за критерієм частотності.

Отримані за схемою $tf\text{-}idf$ ваги g_i нормалізують до 1:

$$g_i = g_i / \sqrt{\sum_{i=1}^h g_i^2}, \quad (5.2)$$

де h – розмір словника.

Нормалізація (5.2) необхідна, тому що невеликі документи мають тенденцію бути представленими короткими векторами термів, в той час як об'ємні документи – довгими. У результаті великі документи мають більш високу ймовірність бути обраними в порівнянні з документами незначного обсягу. Як підсумок (5.1) – (5.2) документ інтерпретується вектором $\{g_i\}_{i=1}^h$ розміром в число термінів словника.

Відзначимо важливі для розпізнавання множин дескрипторів зображень властивості компоненти вектора $tf\text{-}idf$:

- 1) g_i досягає максимуму, якщо термін зустрічається багато раз у невеликій кількості документів (посилуючи їх відміну від інших);
- 2) g_i зменшується, якщо термін зустрічається в якомусь документі лише кілька разів або зустрічається в багатьох документах (тим самим формує менш виражений сигнал про релевантність);
- 3) g_i досягає мінімального значення, якщо термін зустрічається в переважній кількості документів.

5.2. Лінгвістичний аналіз структурних описів

Розглянемо базу зображень у виді скінченної множини $Z = \{Z^i\}_{i=1}^J$ із J описів. Обозначимо потужності множин $s = \text{card } Z$, $s_i = \text{card } Z^i$. При цьому

виконується $s = \sum_{i=1}^J s_i$. Розглянемо опис Z^i як скінченну множину із s_i векторів-дескрипторів SURF розмірністю 64 [25]. Визначимо словник H як підмножину $H \subseteq Z$. Побудуємо словник у базі $Z = \{Z^i\}_{i=1}^J$ шляхом послідовності дій.

1. $\forall z \in Z$ шукаємо подібні їм $z^* \in Z$ за умовою

$$\rho(z, z^*) \leq \varepsilon, \quad (5.3)$$

де ε – поріг подібності; ρ – відстань між дескрипторами SURF у векторному просторі (наприклад, евклідова);

2. Видаляємо із Z елементи, які задовольняють умові (5.3), и вносимо елемент z до словника H , фіксуючи число його повторювань для обчислення (1);

3. Для решти елементів Z виконуємо кроки 1, 2 до тих пір, поки множина Z виявиться порожньою.

У результаті послідовності дій, яка залежить від параметра ε , отримуємо словник $H \subseteq Z$ об'ємом $h = \text{card } H$ елементів. Використовуючи словник H для оброблення кожного із описів Z^i , обчисленням (5.1), (5.2) отримаємо вектор

$$g^i[Z^i] = (g_1, g_2, \dots, g_k, \dots, g_h). \quad (5.4)$$

У результаті застосування лінгвістичного підходу отримана нова форма опису, де еталон Z^i представляється вектором (5.4) із h елементів: $Z^i \rightarrow g^i$.

На етапі розпізнавання об'єкта виконуємо наступні дії. Обчислюємо векторне подання g для розпізнаваного зображення і послідовно знаходимо його відстані до кожного із еталонних векторів. У результаті отримаємо

вектор релевантності $\beta_j = \rho(g, g^j)$, $j=1, \overline{J}$. Об'єкт віднесемо до класу j у відповідності з правилом

$$j = \arg \min_{d=1, \dots, J} \beta_d. \quad (5.5)$$

Схему розпізнавання представимо на рис. 5.1.

Поряд із евклідовою відстанню в (5.5) застосовні також коефіцієнт кореляції, квадрат евклідової відстані, зважена евклідова відстань, хемінгова відстань, відстань Чебишева та ін. міри релевантності [28].

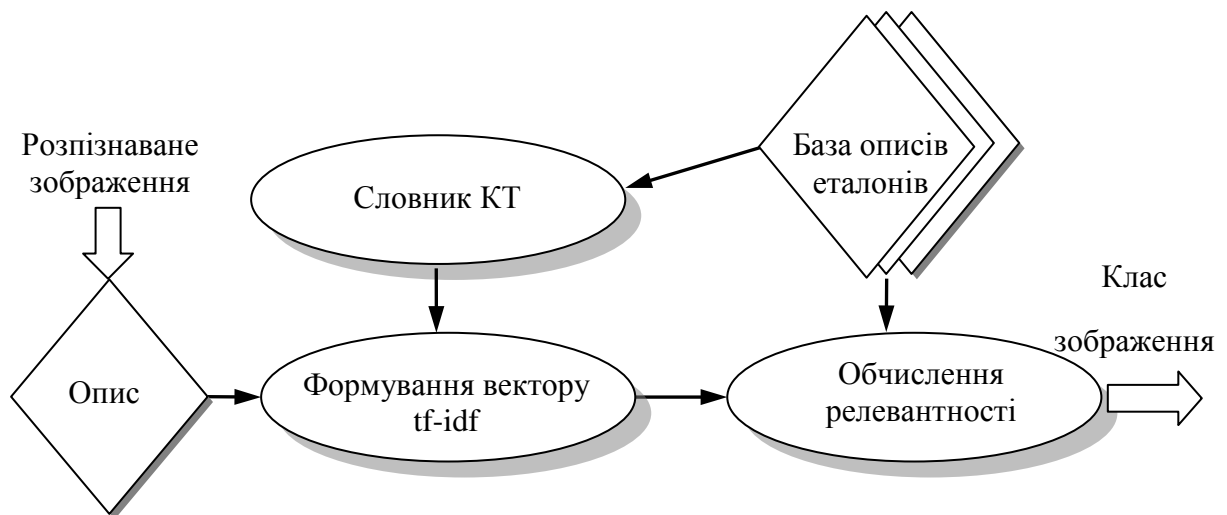


Рис. 5.1 Схема розпізнавання на основі вектору tf-idf я распознавания???

5.3. Застосування медіани для трансформації і зіставлення описів

Узагальнена медіана m для скінченної множини $X = \{x(i)\}_{i=1}^s$ визначається через мінімізацію функціоналу D :

$$m = \arg \min_{v \in X} D(v), \quad D(v) = \sum_{x(i) \in X} \rho[x(i), v], \quad (5.6)$$

де $\rho[x(i), \nu]$ – відстань між елементами множини [53]. Значення D тим менше, чим менше відрізняється число елементів справа і зліва від значення m за відстанню до m . Визначимо медіану m^i для еталону $Z^i = \{z_1, z_2, \dots, z_{s_i}\}$ процедурою:

1) обчислимо симетричну матрицю $R = \{r_{ud}\}$ відстаней розміром $s_i \times s_i$ між елементами Z^i ; на діагоналі матриці R стоять нулі;

2) визначимо суми D_1, D_2, \dots, D_{s_i} елементів у рядках чи стовпцях матриці $D_u = \sum_{d=1}^{s_i} r_{ud}$;

3) обчислимо $m = \arg \min_{\nu \in Z^i} D(\nu)$.

У результаті визначаємо медіану m^i як один із елементів $m^i \in Z^i$. З метою подальшого застосування відсортуємо елементи структурного опису z_1, z_2, \dots, z_{s_i} за зростання значень сум $D_1 \leq D_2 \leq \dots \leq D_{s_i}$. Назвемо цю дію «медіанна обробка опису» [25]. У відсортованому списку $z^* = z_1^*, z_2^*, \dots, z_{s_i}^*$ медіана m^i множини Z^i буде розміщена на першому місці. Якщо замість метрики $\rho[x(i), \nu]$ при обчисленні медіани використовувати міру подібності, наприклад, нормований коефіцієнт кореляції векторів

$$\gamma[x(i), \nu] = \frac{(x(i), \nu)}{\|x(i)\| \cdot \|\nu\|}, \quad (5.7)$$

де в чисельнику стоїть скалярний добуток, а в знаменнику – добуток норм векторів $x(i), \nu$, то медіана відповідає максимуму критерію (5.7), а медіанна обробка призведе до ранжованого за зменшенням списку дескрипторів (7): $D_1 > D_2 > \dots > D_{s_i}$.

Виникає питання, наскільки інформативною є медіана структурного опису в плані розрізнення еталонів Z^i між собою? З іншого боку, цікаво, які вигоди можна отримати в результаті здійснення медіанної обробки?

На наш погляд, у загальному випадку на перше питання однозначної відповіді або оцінок не існує. Наприклад, якщо еталони складаються з різних, але однакових всередині еталона дескрипторів, то навіть по одному представнику (це точно буде медіана) їх можна впевнено розрізнити між собою. Але, як показують наші дослідження [28], описи довільних об'єктів найчастіше складаються із груп досить близьких між собою елементів. Тому для одних об'єктів медіани можуть значно відрізнитися, в той час як для інших ситуацій вони виявляються досить близькими. Погодимося, що в загальному випадку надійно розрізнити два множини по одному їх представнику досить складно, хоча сама медіана вже є результатом деякого інтегрального аналізу множини дескрипторів. Більш універсальним і ефективним способом нам представляється побудова за результатами медіанної обробки списку елементів, який більш точно відображає індивідуальні відмінні властивості всієї множини.

З огляду на те, що принцип побудови медіани призводить до строгої фіксації послідовності елементів у трансформованій множині дескрипторів, з'являється унікальна можливість переробити вихідний опис до стисненого списку, де кожен елемент-дескриптор представлений окремою компонентою вектора. У результаті визначення подібності двох описів можна буде звести до обчислення подібності двох масивів векторів, що, безсумнівно, більш ефективно в прикладному і обчислювальному аспекті, ніж зіставлення скорочених підмножин елементів описів.

Обговоримо можливі шляхи побудови міри подібності (релевантності) описів на основі отриманого в результаті медіанного оброблення списку дескрипторів $z^* = z_1^*, z_2^*, \dots, z_{s_j}^*$. Розглянемо два списки $z^*(i)$, $z^*(k)$ дескрипторів для описів різних об'єктів, що мають в загальному випадку різні обсяги.

Запропонуємо варіанти побудови міри подібності списків:

1) обчислимо подібність як відстань або коефіцієнт кореляції (5.7) між їх медіанами, тобто між першими елементами списків;

2) за встановленим правилом $H: z^* \rightarrow z_h^*$ виберемо із кожного списку $z^*(i), z^*(k)$ фіксоване число q елементів та складемо скорочені списки $z_h^*(i), z_h^*(k)$ як масиви із однакового числа q елементів.

Наприклад, до підписку z_h^* із $q=3$ елементів можна включити перший, середній і останній елементи списку z^* ; до підписку із $q=10$ елементів можна включити перші і останні 3 елементи з кожного списку, а також центральні 4 елементи. Тепер міру подібності μ списків визначимо як

$$\mu[z^*(i), z^*(k)] = \sum_{\nu=1}^q \psi(z_h^*(i, \nu), z_h^*(k, \nu)), \quad (5.8)$$

тобто як суму подоб елементів сформованих підписків; в якості прикладу ψ в (5.8) можна використати (5.7) або евклідову відстань.

Традиційна міра релевантності, в основі якої лежить голосування елементів еталону, відображає значення числа голосів еквівалентних елементів об'єкта щодо максимально можливого числа голосів еталона, тобто

$$\mu[Z^i, Z^*] = n^* / s_i, \quad (5.9)$$

де n^* – число елементів об'єкта Z^* , визнаних еквівалентними елементам еталона в плані виконання умови

$$\psi(z(i), z^*) \leq \varepsilon_z, \quad z(i) \in Z^i, z^* \in Z, \quad (5.10)$$

а значення порогу ε_z для $\psi(\dots)$ встановлює рівнозначність елементів. При обчисленні коефіцієнта (5.7) нерівність в (5.10) змінює вид: $\gamma(z(i), z^*) > \varepsilon_z$. Значення ε_z являється ключовим параметром для процедур голосування [28].

5.4. Аналіз результатів комп'ютерного моделювання

Для проведення експериментів розроблено програмний модуль на мові Java. Обчислення дескрипторів КТ методом SURF здійснено із застосуванням бібліотеки комп'ютерного зору OpenCV для бази із $b=5$ зображень назв брендів фірм (рис. 5.2). На рис. 5.3 у вигляді квадратиків показані координати множини КТ для зображення рис. 5.2, д, а на рис. 5.4 – приклад зображення, отриманого масштабуванням і поворотом еталонного зображення бренду Lenovo. Таблиця 5.1 містить кількості КТ в описах зображень рис. 5.2.

Загальна кількість КТ із бази брендів дорівнювала 1999. Для формування словника в (3) експериментально обраний поріг $\varepsilon=0,6$. Величина порога визначена з міркування, щоб розподіли дескрипторів словника по частоті були приблизно однаковими. За рахунок цих компромісних дій розмір словника буде невеликим, що скорочує обчислення, і в той же час не дуже маленьким, щоб не погіршити якість розпізнавання.



Рис. 5.2 Зображення брендів з бази (а-д).



Рис. 5.3 Координати КТ для зображення рис. 5.2, д



Рис. 5.4 Аналізоване зображення

Таблиця 5.1. Кількості КТ в описах бази зображень рис. 5.2.

Бренди	Acer	Asus	Dell	HP	Lenovo
Число КТ	237	352	638	405	367

У результаті оброблення отриманий словник обсягом 42 КТ, застосовуючи який обчислені і нормалізовані вектора tf-idf еталонів, матриця евклідових відстаней між векторами представлена в таблиці 5.2.

Таблиця 5.2. Матриця відстаней між еталонами

Бренды	Acer	Asus	Dell	HP	Lenovo
Acer	0	0,8579	0,6952	1,3414	0,7991
Asus	0,8579	0	1,0965	1,3514	0,8126
Dell	0,6752	1,0965	0	1,3961	1,0315
HP	1,3414	1,3514	1,3961	0	1,3448
Lenovo	0,7991	0,8126	1,0315	1,3448	0

Як бачимо із таблиці 5.2, у просторі tf-idf спостерігаються істотні відмінності між структурними описами, що свідчить про високий ступінь розрізнення еталонів бази. Використовуючи побудований словник, обчислюємо релевантність між розпізнаваним зображенням рис. 5.4 і еталонами. Результат представлений в таблиці 5.3.

Таблиця 5.3. Значення релевантності між зображенням і еталонами

Бренди	Acer	Asus	Dell	HP	Lenovo
Зображення	0,8372	0,8087	1,0613	1,2151	0,2977

Як бачимо із таблиці 5.3, мінімальна релевантність відповідає еталону бренду Lenovo. Його значення 0,2977 значно менше, ніж найближча за величиною відстань 0,8087, що підтверджує високу якість розпізнавання із застосуванням лінгвістичного підходу. Наше моделювання показало, що обсяг комп'ютерного часу для обчислення релевантності двох описів із застосуванням лінгвістичного уявлення приблизно в 10 разів менший, ніж шляхом застосування традиційної міри (5.9).

Тепер вивчимо питання ефективності розпізнавання і оцінювання швидкодії оброблення із застосуванням запропонованої модифікації релевантності на базі медіанної обробки. За базу зображень використаний набір із 6 зображень чоловічків із заставки Google, приклад зображення і координати точок його структурного опису показаний на рис. 5.5. На рис. 5.6 розміщений вектор медіани для опису зображення рис. 5.5.



Рис. 5.5. Приклад зображення і його множини координат КТ

В ході експериментів у порівняльному плані досліджувався ряд мір подібності структурних описів еталона і розпізнаваного зображення:

- 1) голосування множини дескрипторів КТ і обчислення показника (5.9);
- 2) модифікації (5.8) на основі медіанної обробки при різних значеннях параметра q : $q=1$ – на основі тільки медіани; $q=3$ – перший, середній і останній дескриптор ранжированного опису за критерієм (6); $q=10$, $q=30$ – фіксована кількість дескрипторів, відібраних відповідно до обраної процедури стиснення після медіанної обробки.

Для обчислення подібності дескрипторів у експерименті використовувався коефіцієнт кореляції. Число КТ в описах 6-ти еталонів-чоловічків, один з яких показаний на рис. 5.5, склало: 198, 206, 183, 184, 174, 173.

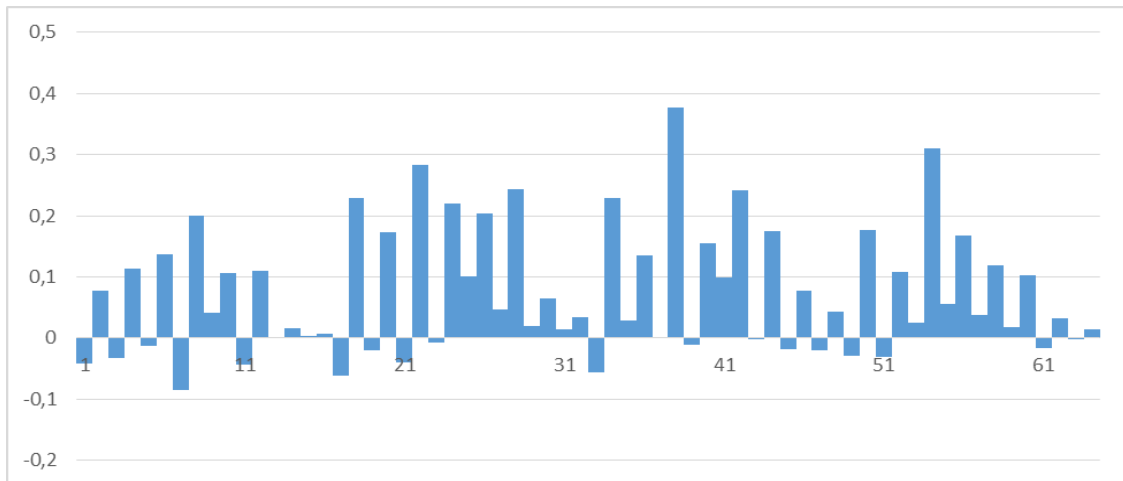


Рис. 5.6. Вектор медіани для опису рис. 5.5

Обчислені матриці релевантності розміром 6×6 для бази еталонів для розглянутих варіантів обробки, де на діагоналі стоять 1. Зрозуміло, що числові оцінки подібності безпосередньо залежать від змісту бази зображень, тому що рівень результативності цілком визначається набором розпізнаваних об'єктів. Максимальне значення недиагональних елементів обчисленої матриці (оцінка рівня помилкового збігу) склало: для голосування - 0,67 (еталони 1,2), для модифікацій $q=1$ - 0,93 (еталони 4,5), $q=3$ - 0,69 (еталони 3,4), $q=10$ - 0,7 (еталони 2,3), $q=30$ - 0,73 (еталони 2,3).

Значення показника помилкової подібності природно залежить від застосовуваної міри, характеристика модифікованого методу при $q=3$ практично збігається з показником для голосування, а модифікації при $q=10$, $q=30$ практично не покращують величину показника помилкового збігу. Навіть при поданні опису тільки медіаною ($q=1$) об'єкти досліджуваної бази все-таки різняться між собою, хоча і в меншій мірі. Для деяких варіантів списку дескрипторів, наприклад, при рівномірному виборі, максимум серед недиагональних елементів матриці подібності досягає 0,85, забезпечуючи впевнене розпізнавання.

Вивчимо тепер питання швидкодії розпізнавання з використанням засобів медіанної обробки. Табл. 5.4 містить оцінки часу обчислення матриці подібності для різних варіантів формування списку (мс).

Таблиця 5.4. Оцінка часу обчислення для варіантів обробки

Метод оброблення	Голосування (9)	$q=1$	$q=3$	$q=10$	$q=30$
Час (мс)	12029	10	12	13	41

Як бачимо із таблиці 5.4, застосування медіанного аналізу за рахунок переходу у простір векторів вражаюче скорочує час оброблення (у 1200 разів у порівнянні із голосуванням!), у той час як показники результативності розпізнавання, які оцінюються матрицею подібності, в прикладній базі зображень практично відповідають традиційному методу. Такий значний виграш у часі можна пояснити суттєвим спрощенням процесу оброблення. Виграш у швидкодії зростає зі збільшенням числа еталонів і зменшується із ростом q . Судячи із досліджуваних варіантів обробки, найкращою в плані показника часу і ефективності є модифікація медіанної обробки при $q=3$.

Результативність медіанної обробки також підтвердилася в ході експерименту, де вхідне зображення піддавалося геометричним перетворенням повороту і масштабування, а потім обчислювалася його релевантність із еталоном для різних модифікацій. Якщо для традиційного методу голосування релевантність знизилася з 1 до 0,7, то для обговорюваних модифікацій значення становлять: при $q=1$ – 0,9, $q=3$ – 0,53. Як бачимо, навіть використання лише однієї медіани забезпечує достатній рівень інваріантності до геометричних перетворень і є кращим за показники традиційного підходу.

5.5. Обговорення досліджень модифікацій методів

Структурні методи розпізнавання зображень на підставі формування опису об'єктів у вигляді множини КТ набули застосування завдяки таким практично важливим властивостям, як інваріантність отриманого опису до геометричних перетворень об'єктів, можливість прийняття рішення за неповним описом, стійкість до завад та ін. [31, 32]. Основні проблеми впровадження цих методів у прикладних системах комп'ютерного зору пов'язані з вдалим вибором типу дескриптора КТ, встановлення необхідної для успішного розпізнавання кількості КТ у описах, поданні опису у стисненому вигляді задля скорочення обчислювальних затрат, пов'язаних із процедурами визначення релевантності для множин КТ, що описують зображення об'єктів.

Популярність детекторів BRISK пов'язана з поданням дескриптора КТ у виді бінарного вектору ознак, що сприяє можливості застосування обговорюваних методів у мобільних пристроях. Бінарний образ дескриптора КТ не тільки спрощує та прискорює обчислення, але й досить ґрунтовно відображає властивості фрагментів зображення, що в подальшому сприяє якісному розпізнаванню.

Напрямки сучасних досліджень, що в цілому сприяють підвищенню ефективності методів, пов'язані зокрема з такими задачами, як подання множини дескрипторів у кластерному виді [2, 12, 31, 32], формуванні «спільного чи мега- дескриптора» на підставі множини дескрипторів КТ усього зображення, застосуванні хешування у просторі дескрипторів [18, 19, 31].

Тут наведено результати проведення експериментального дослідження модифікацій методу класифікації зображень з використанням подібності описів у просторі дескрипторів на підставі впровадження апарату аналізу бітових даних.

Опишемо та дослідимо експериментально наступні модифікації методу при здійсненні зіставлення описів задля визначення їх релевантності [31].

Традиційний метод для обчислення релевантності двох описів O_1, O_2 зводиться до підрахунку числа r «еквівалентних» елементів одного з описів, що вважається еталоном, у складі іншого опису [28]

$$r = \text{card}\{x_1 \in O_1 \mid x_1 \sim x_2, x_2 \in O_2\}. \quad (5.11)$$

Еквівалентність $x_1 \sim x_2$ в (5.11) трактується як

$$x_1 \sim x_2 : \rho(x_1, x_2) \leq \delta, \quad (5.12)$$

де ρ – деяка метрика у просторі B^n , δ – поріг для значення метрики, в межах якого елементи x_1, x_2 вважаються еквівалентними.

Як правило, прийняття рішення ґрунтується на нормованому значенні r , де параметром нормування є чисельність $m_1 = \text{card}\{O_1\}$ елементів етальної множини O_1 . Для дескрипторів КТ з простору B^n у виразі (5.12) можна використати нескладну для обчислень метрику Хемінга. Більш проблематичні за обсягом обчислень підходи пов'язані з визначенням мір подібності чи відстані у просторі описів як множин КТ [28].

Перспективною ідеєю задля суттєвого скорочення об'єму обчислень є кластерне подання [2, 12], в результаті чого структурний опис подається у вигляді

$$h[O] = (h_1, h_2, \dots, h_k),$$

де k – число сформованих кластерів, h_i – число елементів опису, що віднесені у процесі кластеризації до кластеру з номером k . Як правило,

квантування здійснюють в межах бази зображень, тобто на фіксованій множині еталонних описів.

У результаті такої трансформації опис представляється вектором, при визначенні релевантності описів обчислюється подібність векторів h (а не множин O векторів). Як показують експерименти, час розпізнавання практично зі збереженням якості навіть для невеликого числа еталонів (5-10) скорочується в десятки разів [2]. Параметрами такої модифікації є вибраний метод кластеризації, кількість кластерів, значення порогу для встановлення еквівалентності дескрипторів та ін.

Узагальнення ідеї кластерного квантування простору дескрипторів КТ приводить до думки на основі опису еталона як множини векторів побудувати деякий узагальнений дескриптор $d[O] = (d_1, d_2, \dots, d_n)$, $d \in B^n$, що відображає властивості всієї множини O . Задача побудови дескриптора $d[Z^j]$ для еталону Z^j класу j реалізується шляхом відображення $D: Z^j \rightarrow d[Z^j]$, результатом якого є вектор $d \in B^n$. Відображення D трансформує множину векторів у вектор, що належить B^n .

Побудова спільного дескриптора націлена не тільки на скорочення обсягу обчислювальних витрат, стільки на узагальнення і концентрування важливої для класифікації інформації про розпізнавані об'єкти в аспекті побудови ефективного класифікаційного правила. Результат являє собою деякий узагальнений «центр» еталонного опису. Важливою властивістю узагальненого дескриптора є вимога його незалежності від порядку слідування дескрипторів у описові, так як це гарантує інваріантність його значень відносно групи геометричних перетворень, до яких інваріантні дескриптори [4, 12, 19].

Одним із найпривабливіших механізмів для побудови «спільного дескриптора», що виступає як мега-центр даних, є апарат аналізу бітів [4]. Наприклад, спільний дескриптор d можна побудувати шляхом інтелектуального аналізу та узагальнення значень груп відповідних бітів для

кожного з елементів опису. Результируючий дескриптор d при цьому може бути складено з комбінацій бітів, що найчастіше зустрічаються у множині елементів $x \in O$, тобто мають найбільшу підтримку.

У межах апарату бінарних даних викликає інтерес застосування націленого оброблення множини дескрипторів із визначенням одного з елементів, що буде представляти множину як клас. Двійковий вид дескрипторів дає підвалини застосувати апарат логічного оброблення. Зважаючи на бінарний вид BRISK- та ORB-дескрипторів, для кожного з еталонів Z^j визначимо бінарний вектор m_j центра класу на підставі такого логічного правила [4, 19]

$$m_j(a) = \begin{cases} 1, & f(Z^j, a) \geq s(j) / 2, \\ 0, & \text{інакше,} \end{cases} \quad (5.13)$$

де $m_j(a)$ – значення біту на позиції з номером a для центру m_j , $f(Z^j, a)$ – функція, що підраховує кількість одиничних бітів на позиції з номером a у множині дескрипторів еталону Z^j , $a = 1, \dots, n$. Функція $f(Z^j, a)$ може бути обчислена безпосередньо додаванням бітів як

$$f(Z^j, a) = \sum_{d=1}^{s(j)} x_d(a), \quad x_d \in Z^j, \quad (5.14)$$

де $x_d(a)$ – біт з номером a для дескриптора з номером d в описі еталона.

За виразом (5.13) значення кожного з бітів центру m_j для j -го класу визначається переважною більшістю значень відповідних розрядів усіх дескрипторів КТ, які належать еталону Z^j . Бітове подання можна ефективно використати при роботі з множинами даних. Підрахунок числа одиничних

бітів у (5.14) можна здійснити також застосуванням відповідних бінарних функцій.

Аналізуємо елементи $x \in Z$ загального вмісту структурних описів бази еталонів (навчальна вибірка) шляхом віднесення їх до відповідного класу з використанням конкуренції. У якості ρ застосуємо метрику Хемінга

$$\rho(x, m_j) = \sum_{a=1}^n |x(a) - m_j(a)|, \quad (5.15)$$

що визначає кількість розбіжних бітів для двійкових послідовностей однакової довжини.

У результаті оброблення з використанням попередньо обчислених центрів одержуємо подання для кожного еталона: $h[Z^i] = (h_1^i, \dots, h_J^i)$, яке відповідає розподілу елементів множини Z^i за класами еталонів, де h_a^i – цілі числа, причому $\sum_{a=1}^{s(i)} h_a^i = s(i)$, так як кожний з елементів опису буде віднесено до одного з фіксованих класів.

Оброблення за аналізом значень бітів або байтів формує значимі властивості сукупності елементів опису, а послідовне об'єднання їх у новий вектор зберігає вихідну структуру даних і їх уніфікацію. Найбільш детальний аналіз зводиться до побітового.

Дієвим засобом підвищення швидкодії оброблення об'ємних масивів багатовимірних даних є апарат хешування. Суть хешування полягає в тому, що потенційно нескінченне число оброблюваних елементів представляється у вигляді скінченної множини їх класів. Як правило, його застосовують для задач пошуку, сортування [10]. Практика показує, що ефективним є також використання хеш-функцій і при зіставленні масивів даних, в тому числі і для множин дескрипторів КТ [10]. Зважаючи на те, що дескриптори BRISK є бінарними, для визначення хеш-функцій можна застосувати аналіз значень комбінацій бітів.

Нехай b – задане число класів для хешування, а $g(x)$, $x \in V^n$ – хеш-функція $g(x) \in [0, \dots, b-1]$, що задана на інтервалі значень цілих чисел. Результатом оброблення опису O буде хеш-таблиця, тобто масив із b списків, що містять групи дескрипторів із фіксованими значеннями $g(x)$. Результативність такого подання для прикладної задачі безпосередньо залежить від виду вибраної хеш-функції, кількості b класів, а також безпосередньо від оброблюваних даних. Один із поширених видів хеш-функції є значення кількості одиничних бітів у бінарному коді.

Порівняння двох хеш-таблиць з метою отримання ступеня їх релевантності – це окреме важливе завдання. Воно може бути реалізоване у нашій постановці як відстань у цілочисельному векторному просторі між векторами, що містять кількості елементів даних з фіксованим значенням $g(x)$. Інший спосіб полягає у реалізації підрахунку кількості елементів із однієї хеш-таблиці, що входять у склад другої. У будь-якому випадку хешування забезпечує суттєве скорочення обсягу обчислень за рахунок спеціальної організації пам'яті [10].

Програмне моделювання досліджуваних модифікацій методів створено мовою C# у середовищі Visual Studio 2017 з використанням засобів бібліотеки Open CV [86].

Оцінювання результативності застосування кластерного апарату відбувалось на прикладі задачі ідентифікації людських облич, зокрема, з метою виявлення ступеня схожості фотографій близнюків та родичів. На рис. 5.7, 5.8 наведено приклади досліджуваних об'єктів за зображеннями розміром 250x260. Експерименти проводилися з числом кластерів $k=2, 3, 5$. Приклад координат виділених детектором BRISK КТ представлено на рис. 5.9. Кількість встановлених КТ для зображень рис. 5.7 дорівнювала 160, для зображень рис. 5.8 – 115.



Рисунок 5.7. Зображення близнюків



Рисунок 5.8. Зображення матері та доньки



Рисунок 5.9. Зображення з координатами КТ

Кластеризація в межах бази із двох зображень виконана методом к-середніх. Особливістю є бінарний вид множини дескрипторів. Модифікація методу кластеризації зводилась до того, що на кожному кроці новий центроїд кластера формується так: у поточній множині елементів кожного кластера побітово рахується кількість 0 та 1, якщо кількість 1 більше, то біт центроїда встановлюється рівним 1, навпаки – 0 [4, 10].

Особливістю результатів кластеризації для обох типів зображень є зосередження їх кластерних подань фактично у 2-х кластерах. Для рис. 5.7 пропорції числа КТ склали приблизно 90-60, а для рис. 5.8 – 93-22 незалежно від числа кластерів. Така практична ситуація, як правило, сприяє кращому розрізненню об'єктів.

Значення манхеттенської відстані, за якою визначено релевантність для векторних подань двох зображень рис. 5.7 при $k=5$ має значення 7, а для рис. 5.8 – значення 12. Зважаючи, що максимальне значення відстані тут складає 160, а мінімальне – 0, можна зробити незаперечний висновок про достатню адекватність використання кластерної моделі в задачі зіставлення описів таких об'єктів. Загалом відстань для близнюків в ряді інших експериментів була меншою за відстань для матері та доньки для всіх значень k . В той же час відстань у кластерному поданні між представниками зображень рис. 5.7 та рис. 5.8 перевищує значення 130. Це підкреслює чутливість методу до незначних особливостей зображення і його можливість широкого застосування у системах комп'ютерного зору.

Аналіз комп'ютерних витрат часу на обчислення значення релевантності за кластерним поданням у порівнянні з традиційним методом голосування КТ шляхом підрахунку числа еквівалентних дескрипторів для зображень рис. 5.7, 5.8 показав вигреш у 2,5 рази. Вигреш пропорційно зростає зі збільшенням кількості КТ та числа еталонів у задачі класифікації.

Для моделювання процедури зіставлення описів із застосуванням хешування використано зображення тварин (бегемот) розміром 530×270 , кількість КТ детектора BRISK склала величину 960. Здійснено поворот зображення на 15 градусів, обчислена релевантність вхідного та оберненого зображень шляхом голосування значень дескрипторів КТ, а також визначення манхеттенської відстані між побудованими хеш-таблицями. Хеш-функція при $b=512$ підраховувала число одиничних бітів. На рис. 5.10 представлено ненульовий відрізок хеш-таблиці вхідного зображення.

За результатом моделювання релевантність зображень традиційним голосуванням склала величину 0,77 (час – приблизно 27 сек), а шляхом обчислення нормованої відстані між хеш-таблицями – 0,21 (час – 0,033 сек).

Як бачимо, хеш-подання без втрати даних суттєво прискорює процес обчислення ступеня релевантності описів, не погіршуючи значення подібності. Час оброблення для даного прикладу скорочується у порівнянні з традиційним голосуванням приблизно 800 разів і залежить від кількості КТ в описі та від способу визначення релевантності хеш-таблиць.

Реалізація голосування для структури даних у вигляді хеш-таблиць для множин КТ показала, що у цьому випадку результат обчислення релевантності суттєво залежить від вибраної хеш-функції, а сама процедура повинна бути безпосередньо пов'язана з її виглядом.

Експериментальне визначення узагальненого дескриптора здійснювалося на зображеннях ікон (рис. 5.11) двома способами: побітове оброблення та оброблення за парами бітів із внесенням до узагальненого дескриптора значень, що найчастіше зустрічаються у фіксованому структурному описі.

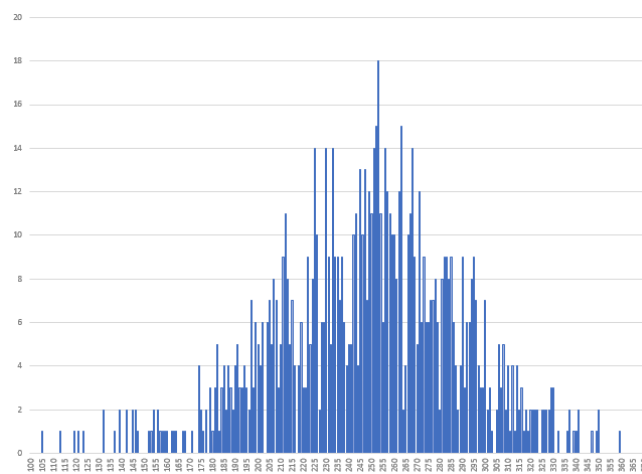


Рис. 5.10. Частина хеш-таблиці зображення



Рис. 5.11. Зображення ікони «Сікстинська мадонна» з координатами КТ

Кількість КТ для зображень ікон 400x540 склала 1350 і скорочувалася в експерименті в цілях дослідження. Релевантність описів обчислювалася як відстань Хемінга між узагальненими дескрипторами та як нормоване число голосів еквівалентних значень дескрипторів КТ для двох різних описів.

Як показало проведене нами моделювання, зі збільшенням числа КТ в описові з використанням вказаних способів побудови узагальненого дескриптора за критерієм частоти входження до опису відстань між узагальненими дескрипторами різних зображень зменшується, тобто ймовірність їх розрізнення дещо погіршується. Так, міра подібності за голосуванням для двох різних ікон склала 0,12 (максимум 1,0), в той час як нормована до числа точок еталону відстань Хемінга для їх узагальнених дескрипторів склала 0,068 (мінімум 0). Але уже для скороченого випадковим чином опису із 50 КТ ця відстань зросла до 0,32. Цей факт можна пояснити зростанням ступеня різноманіття даних всередині опису, що для вибраного виду оброблення призводить до деякого «розмазування» образу у просторі, зменшення рівня індивідуальності його подання та до збільшення рівня перетину образів різних зображень. Для досягнення достатнього ступеня

розрізненості образів тут можна рекомендувати попереднє оброблення опису для включення в його склад тільки найбільш значущих для даного зображення дескрипторів. Для конкретних зображень в експерименті при вказаному способі побудови узагальненого дескриптора рекомендується використовувати 50-70 КТ. Час обчислень для модифікованого методу у порівнянні з голосуванням зменшився приблизно у 10 разів. Для методу попарного аналізу бітів відстань у просторі узагальнених дескрипторів приблизно вдвічі менша, що потребує ще більшого зменшення опису для досягнення чіткого розрізнення за цим способом.

Висновки

1. Основним результатом досліджень цього розділу є побудова і результативне практичне застосування моделей аналізу та оброблення описів: лінгвістичного підходу, медіанної обробки, кластеризації, хешування та визначення мега-центру в завданнях структурного розпізнавання зображень.

2. Перехід до стисненого векторного чи спискового уявлення множини дескрипторів істотно підвищує швидкодію розпізнавання за рахунок спрощення процесу оброблення.

3. Хешування опису без втрати даних суттєво (у експерименті у сотні разів) прискорює процес обчислення релевантності описів. Вибрана хеш-функція може впливати на результат і сприяти покращенню розрізнення зображень.

4. Застосування узагальненої мега-характеристики – спільного дескриптора значно (у десятки разів) скорочує час обчислень, при цьому виникає потреба у попередньому обробленні з метою формування стисненого опису із значущих для таких зображень дескрипторів.

5. Трансформація описів до кластерного подання не тільки скорочує час оброблення, але й демонструє чутливість представлених модифікацій оброблення до характерних особливостей зображення.

Список літературних джерел

1. Гадецька С.В., Гороховатський В.О. (2018) Методи структурної класифікації зображень на засадах баєсовської теорії прийняття рішень. *Радіоелектроніка, інформатика, управління*, №2 (45), с. 90–97.
2. Gorokhovatskyi Volodymyr, Putyatin Yevgeniy, Gorokhovatskyi Oleksii, Peredrii Olena. Quantization of the Space of Structural Image Features as a Way to Increase Recognition Performance. *The Second IEEE International Conference on DataStream Mining & Processing 21-25 August 2018, Lviv, Ukraine*, p. 464 – 467.
3. Гороховатський В.О., Пупченко Д.В., Солодченко К.Г. (2018) Аналіз властивостей, характеристик та результатів застосування новітніх детекторів для визначення особливих точок зображення. *Системи управління, навігації та зв'язку*, №1 (47), с. 93–98.
4. Гороховатський В.О., Солодченко К.Г. (2018) Застосування апарату аналізу та оброблення бітових даних у методах класифікації зображень за множиною ключових точок. *Системи управління, навігації та зв'язку*, №2 (48), с. 63–67.
5. Gorokhovatskyi V.A. (2018) Image Classification Methods in the Space of Descriptions in the Form of a Set of the Key Point Descriptors. *Telecommunications and Radio Engineering*, 77 (9), pp. 787-797.
6. Гадецька С.В., Гороховатський В.О. (2018) Застосування статистичних мір релевантності для векторних структурних описів об'єктів у задачі класифікації зображень. *Системи управління, навігації та зв'язку*, №4 (50), с. 62–68.
7. Gorokhovatskyi Volodymyr, Gadetska Svitlana (2018). Classification of Images of Visual Objects Based on Statistical Relevance Measures of Their Structural Descriptions. *Proc. the IEEE 13th International Scientific and Technical*

Conference on Computer Sciences and Information Technologies (CSIT'2018), 11-14 September, Lviv, Ukraine, pp. 68 – 71.

8. Гадецька С.В., Гороховатський В.О. (2018). Статистичні міри для визначення релевантності структурних описів зображень. Інтелектуальні системи прийняття рішень та проблеми обчислювального інтелекту (ISDMCI'2018): збірка наук. праць міжн. наук. конф., с. Залізний Порт, 21–27 травня 2018. – Херсон: ФОП Вишемирський В.С., с. 224–226.

9. Gadetska S.V., Gorokhovatsky V.O. (2018) Statistical Measures for Computation of the Image Relevance of Visual Objects in the Structural Image Classification Methods. Telecommunications and Radio Engineering, vol. 77 (12), pp. 1041–1053.

10. Gorokhovatskiy V.A., Gorokhovatskiy A.V., Peredrii Ye.O. (2018) Hashing of Structural Descriptions at Building of the Class Image Descriptor, Computing of Relevance and Classification of the Visual Objects. Telecommunications and Radio Engineering, vol. 77 (13), pp. 1159–1168.

11. Gorokhovatskiy O., Gorokhovatskiy V., Peredrii O. (2018) Analysis of Application of Cluster Descriptions in Space of Characteristic Image Features. Data, 3(4), 52. – doi: 10.3390/data3040052. Available online: <https://www.mdpi.com/2306-5729/3/4/52>

12. Гороховатський В.О., Гадецька С.В., Пономаренко Р.П. (2018) Статистичні розподіли та ланцюжкове подання даних при визначенні релевантності структурних описів візуальних об'єктів. Системи управління, навігації та зв'язку, №6 (52), с. 87–92.

13. Gorokhovatsky, V.O., Gadetska, S.V., (2019) Determination of Relevance of Visual Object Images by Application of Statistical Analysis of Regarding Fragment Representation of their Descriptions, Telecommunications and Radio Engineering, 78 (3), pp. 211–220, doi: 10.1615/TelecomRadEng.v78.i3.20.

14. Gorokhovatsky V.O., Gadetska S.V. (2019) Spatial and Statistical Analysis and Processing of Fragments Ensemble of Image Structural Description in Recognition

of Visual Objects. *Telecommunications and Radio Engineering*, 78 (14), pp. 1263-1274, doi: 10.1615/TelecomRadEng.v78.i14.40.

15. Гороховатський В.О., Гадецька С.В., Стяглик Н.І. (2019) Вивчення статистичних властивостей моделі блочного подання для множини дескрипторів ключових точок зображень. *Радіоелектроніка, інформатика, управління*, №2, с. 100–107, doi: 10.15588/1607-3274-2019-2-11.

16. Гадецька С.В., Гороховатський В.О., Пономаренко Р.П. Розпізнавання зображень за розподілами фрагментів їх опису. *Інтелектуальні системи прийняття рішень та проблеми обчислювального інтелекту (ISDMCI'2019): матеріали міжн. наук. конф., с. Залізний Порт, 21–25 травня 2019. – Херсон: ФОП Вишемирський В.С., 2019.– С. 35–37.*

17. Gorokhovatskyi V., Gadetska S., Ponomarenko R. (2020) Recognition of Visual Objects Based on Statistical Distributions for Blocks of Structural Description of Image. *Lecture Notes in Computational Intelligence and Decision Making. Proceedings of the XV International Scientific Conference “Intellectual Systems of Decision Making and Problems of Computational Intelligence” (ISDMCI'2019), Ukraine, May 21–25, 2019, pp. 501-512. – Available online: https://rd.springer.com/chapter/10.1007/978-3-030-26474-1_35*

18. Гороховатський, В.О., Гадецька, С.В., Пономаренко, Р.П. (2019) Логічний аналіз та оброблення даних задля класифікації зображень на підставі формування статистичного центру опису. *Системи управління, навігації та зв'язку. – №4 (56). – С. 43–48. – doi: 10.26906/SUNZ.2019.4.043.*

19. Gorokhovatskyi, V.O., Tvoroshenko, I.S., and Peredrii O.O. (2020) Image classification method modification based on model of logic processing of bit description weights vector, *Telecommunications and Radio Engineering*, 1 (79), pp. 59-69. doi: 10.1615/TelecomRadEng.v79.i1.60. Available online: http://www.dl.begellhouse.com/ru/journals/0632a9d54950b268,706f31e606bab3dd_2492ae044887304a.html

20. Гороховатський В.О., Руденко Д.О., Сірик Т.О. (2019) Дослідження системи ієрархічних ознак при блочному поданні опису у складі множини ключових точок зображення. Системи управління, навігації та зв'язку, №2 (54), с. 69–73, doi: 10.26906/SUNZ.2019.1.003.
21. Gorokhovatsky V.A., Gorokhovatsky A. V., Berestovsky A.Ye. (2016) Intellectual Data Processing and Self-Organization of Structural Features at Recognition of Visual Objects. Telecommunications and Radio Engineering, vol. 75, No 2, pp. 155–168.
22. Гороховатский В.А., Путятин Е.П., Столяров В.С. (2017) Исследование результативности структурных методов классификации изображений с применением кластерной модели данных. Радиоэлектроника, информатика, управление, №3 (42), с. 78–85.
23. Гороховатский В.А., Столяров В.С. (2016) Классификация изображений на основе кластерного представления структурных описаний. Бионика интеллекта, №2 (87), с.83–87.
24. Gorokhovatsky V.A. (2016) Efficient Estimation of Visual Object Relevance during Recognition through their Vector Descriptions. Telecommunications and Radio Engineering, vol. 75, No 14, p. 1271–1283.
25. Гороховатский В.А. (2017) Методы определения релевантности изображений на основе медианной обработки структурных описаний. Радиоэлектроника, информатика, управление, №1 (40), с. 100–106.
26. Путятин Е.П., Аверин С.И. (1990) Обработка изображений в робототехнике, Москва, Машиностроение, 320 с.
27. Путятин Є.П., Гороховатський В.О., Матат О.О. (2006) Методи та алгоритми комп'ютерного зору: навч. посіб., Харків, ТОВ «Компанія СМІТ», 236 с.
28. Гороховатский В.А. (2014) Структурный анализ и интеллектуальная обработка данных в компьютерном зрении, Харьков, Компания СМІТ, 316 с.
29. Гороховатский В.А., Полякова Т.В. (2018) Применение пространственных

структур признаков для классификации изображений в компьютерном зрении, Харьков, ФОП Панов А.Н., 120 с.

30. Гороховатский В.А. (2003) Распознавание изображений в условиях неполной информации, Харьков, ХНУРЭ, 112с.

31. Гороховатський В.О., Васильченко А.А., Манько К.П., Пономаренко Р.П. (2018) Дослідження модифікацій методу встановлення релевантності зображень об'єктів за описами у вигляді множини дескрипторів ключових точок. Системи управління, навігації та зв'язку, №5 (51), с. 74–78.

32. Гороховатский В.А., Ересько Ю.Н., Путятин Е.П., Стрельченко В.И. (1990) Локализация объектов на изображениях визуальных сцен. Автометрия, №6, с. 3–7.

33. Putyatin Y., Gorokhovatsky V., Gorokhovatsky A., Peredriy E. (2008) Projective methods of image recognition. In: Intelligent Technologies and Application. K. Markov et al. (Eds.), Sofia: FOI ITHEA, Information Science and Computing, No 5, p. 37-43.

34. Gorokhovatsky V.A., Kobylin O.A., Kulikov Yu.A. (2015) Application of Granulation of Feature Descriptions in Structural Image Recognition. Telecommunications and Radio Engineering, 74 (6), p. 503–514.

35. Гороховатский В.А., Путятин Е.П. (2008) Структурное распознавание изображений на основе моделей голосования признаков характерных точек. Реєстрація, зберігання і обробка даних, т.10, №4, с.75–85.

36. Gorokhovatskiy, V.A. (2011) Compression of Descriptions in the Structural Image Recognition. Telecommunications and Radio Engineering, 70 (15), pp. 1363–1371, doi: 10.1615/TelecomRadEng.v70.i15.60.

37. Bay H., Tuytelaars T., Van Gool L. (2006) Surf: Speeded up robust features. Computer Vision: Ninth European Conference on Computer Vision, Graz, 7-13 May, 2006: proceedings, Berlin, Springer, p.404 – 417.

38. Rublee, E., Rabaud, V., Konolige, K., and Bradski, G., (2011) ORB: an efficient alternative to SIFT or SURF, IEEE International Conference on Computer Vision (ICCV), Proceedings, pp. 2564-2571.
39. Leutenegger, S., Chli, M., and Siegwart, R.Y., (2011) BRISK: Binary Robust Invariant Scalable Keypoints. – Computer Vision (ICCV), pp. 2548-2555.
40. Zadeh L. (1997) Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic. *Fuzzy Sets Syst*, vol. 90, p. 111-127.
41. Шапиро Л., Стокман Дж. (2006) Компьютерное зрение. Москва, БИНОМ. Лаборатория знаний, 752 с.
42. Форсайт Д., Понс Ж. (2004) Компьютерное зрение. Современный подход, Москва, Издательский дом «Вильямс», 928 с.
43. Шлезингер М., Главач В. (2004) Десять лекций по статистическому и структурному распознаванию, Киев, Наукова думка, 535с.
44. Машталир В.П. (2001) Точечно-множественные методы обработки информации, Харьков, Бизнес Информ, 199 с.
45. Гонсалес Р., Вудс Р. (2005) Цифровая обработка изображений, Москва, Техносфера, 1070 с.
46. Bian J, Lin W Y, Matsushita Y et al. (2017) GMS: Grid-Based Motion Statistics for Fast, Ultra-Robust Feature Correspondence. IEEE Conference on Computer Vision and Pattern Recognition (CVPR).
47. Y. Zhang, C. Li, C. Cao et al. (2018) An Improved ORB Feature Point Matching Algorithm. Proceedings of the 2018 2nd International Conference on Computer Science and Artificial Intelligence, pp. 207-211.
48. Calonder, M., Lepetit, V., Strecha, C., & Fua, P. (2010, September). Brief: Binary robust independent elementary features. In European conference on computer vision (pp. 778-792). Springer, Berlin.
49. Szeliski R. (2010) Computer Vision: Algorithms and Applications, London, Springer, 979 p.

50. Manning C.D., Raghavan P., Schütze H., (2008), Introduction to Information Retrieval, Cambridge, University Press.– 528p.
51. Солтон Дж. (1979) Динамические библиотечно-информационные системы, Москва, Мир, 558 с.
52. Han, J. and Kamber, M., (2006) Data Mining: Concepts and Techniques, Amsterdam: Morgan Kaufmann Publishers, 754 p.
53. Кохонен Т. (2013) Самоорганизующиеся карты, Москва, БИНОМ. Лаборатория знаний, 655 с.
54. R. Xu and D. C. Wunsch (2009) Clustering. IEEE Press Series on Computational Intelligence. Hoboken, NJ: John Wiley & Sons, Inc.
55. Aggarwal C.C. Data Mining. Cham: Springer, Int. Publ. Switzerland, 2015.
56. Gorokhovatsky A. V., Gorokhovatsky V.A., Vlasenko A.N., Vlasenko N.V. (2014) Quality Criteria for Multidimensional Object Recognition Based Upon Distance Matrices. Telecommunications and Radio Engineering, vol. 73, №18, p. 1661 – 1670.
57. Shafronenko, A., Bodyanskiy, Ye., Rudenko D.: Neuro-Fuzzy Clustering Of Distorted Data Using Cat Swarm Optimization. LAP LAMBERT Academic Publishing, 2020.
58. Kacprzyk, J., Pedrycz, W. (2015) Springer Handbook of Computational Intelligence, Springer-Verlag, Berlin Heidelberg.
59. Воробель Р.А. (2012) Логарифмічна обробка зображень, Київ, Наукова думка, 232 с.
60. Флах П. (2015) Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных, Москва, ДМК Пресс, 400 с.
61. Gorokhovatsky V. A., Putyatin Ye. P. (2009) Image Likelihood Measures of the Basis of the Set of Conformities. Telecommunications and Radio Engineering, vol. 68 (9), pp. 763-778.

62. Gorokhovatskyi Volodymyr, Tvoroshenko Iryna (2020) Image Classification Based on the Kohonen Network and the Data Space Modification. In CEUR Workshop Proceedings: Computer Modeling and Intelligent Systems (CMIS-2020). 2608. pp. 1013-1026. Available online: <http://ceur-ws.org/Vol-2608/>
63. Gorokhovatskyi Volodymyr, Gadetska Svitlana, Stiahlyk Natalia (2020) Image structural classification technologies based on statistical analysis of descriptions in the form of bit descriptor set. In CEUR Workshop Proceedings: Computer Modeling and Intelligent Systems (CMIS-2020). 2608. pp. 1027-1039. Available online: <http://ceur-ws.org/Vol-2608/>
64. Filatov, V., Kovalenko, A. (2019). Fuzzy Systems in Data Mining Tasks. Studies in Computational Intelligence, 243–274. doi:10.1007/978-3-030-35480-0_6
65. Гороховатский, В.А., Гороховатский, А.В., Берестовский, А.Е. (2016) Структурное распознавание изображений с применением моделей интеллектуальной обработки и самоорганизации признаков. Радиоэлектроника, информатика, управление, №3 (38), с. 39–46.
66. Гороховатский В.А. Куликов Ю.А. (2012) Формализм мультимножеств в задачах структурного распознавания и поиска в базах видеоданных. Искусственный интеллект, №1, с. 76–85.
67. Karami E. Image Matching Using SIFT, SURF, BRIEF and ORB: Performance Comparison for Distorted Images / E. Karami, S. Prasad, M. Shehata. – Режим доступа https://www.researchgate.net/publication/292157133_
68. Duda R.O., Hart P.E., Stork D.G (2000) Pattern classification, 2ed., Wiley, 738p.
69. Вапник В.Н. , Червоненкис А.Я. (1974) Теория распознавания образов (статистические проблемы обучения), Москва, Наука, 416 с.
70. Феллер В. (1984) Введение в теорию вероятностей и ее приложения. Т. 1. Москва, Мир, 528 с.

71. Айвазян, С.А., Енюков, И.С., Мешалкин, Л.Д. (1983). Прикладная статистика: Основы моделирования и первичная обработка данных, Москва, Финансы и статистика, 471 с.
72. Гуров С.И. (2002) Как оценить надежность алгоритма классификации. Таврический вестник информатики и математики, Симферополь, КНЦ НАН Украины, № 1, с. 27 – 56.
73. Фукунага К. (1979) Введение в статистическую теорию распознавания образов, Москва, Наука, 367с.
74. Ампилова Н.Б. Применение расхождения Реньи к анализу и классификации изображений. / Н.Б. Ампилова, В.Д. Сергеев, И.П. Соловьев. [Электронный ресурс]. – Режим доступа:
https://lib.herzen.spb.ru/media/magazines/contents/1/176/ampilova_176_35_44.pdf
75. Худов В.Г., Маковейчук, О.М., Хижняк І.А. (2017) Оцінка якості сегментування оптико-електронного зображення шляхом оцінки комплексних показників на відстані Кульбака-Лейблера. Системи обробки інформації, 4 (150), с. 27 – 30.
76. Radu Stoica. Indexing and Retrieval in Multimedia Libraries Through Parametric Texture Modeling using the 2D Wold Decomposition. [Электронный ресурс]. – Режим доступа:<https://hal.inria.fr/inria00073085/document>.
77. Кульбак С. (1967) Теория информации и статистика, Москва, Наука, 408 с.
78. David J.C. MacKay. Information Theory, Inference, and Learning Algorithms. [Электронный ресурс]. – Режим доступа:
<http://www.inference.org.uk/itprnn/book.pdf>.
79. Бочаров П.П., Печинкин А.В. (2005) Теория вероятностей. Математическая статистика, Москва, ФИЗМАТГИЗ, 296 с.
80. Alfred O. Hero, Bing Ma, Olivier Michel, John Gorman. Alpha-Divergence for Classification, Indexing and Retrieval. [Электронный ресурс]. – Режим доступа:
<https://pdfs.semanticscholar.org/05a8/1e8891f76995b30267be776c78013a9143a3.pdf>

81. Tim van Erven, Peter Harremoës. Rényi Divergence and Majorization. [Електронний ресурс]. – Режим доступу: <https://arxiv.org/pdf/1001.4448.pdf>.
82. Porter F. Testing Consistency of Two Histograms. [Електронний ресурс]. – Режим доступу: https://www.researchgate.net/publication/1917663_Testing_Consistency_of_Two_Histograms.
83. Вадзинский Р. (2008) Статистические вычисления в среде EXCEL. Библиотека пользователя, СПб, Питер, 608 с.
84. Гмурман В.Е. (2004) Теория вероятностей и математическая статистика. Москва, Высш. шк, 479 с.
85. Волькенштейн М.В. (2006) Энтропия и информация. Москва, Наука, 106 с.
86. OpenCV foundation. (2017). OpenCV: Image Thresholding. Retrieved from https://docs.opencv.org/3.3.0/d7/d4d/tutorial_py_thresholding.html
87. Гороховатський В.О., Дубницький В.Ю., Кобилін А.М., Лукін В.О., Москаленко О.В. (2016) Методи оцінки витрат на розробку програмного забезпечення: монографія, Харків, ТОВ «Компанія СМІТ», 2016, 108с.
88. Гороховатский В.А. (2008) Иерархия пространственных отношений структурных признаков в задачах сопоставления визуальных объектов. Системы управління, навігації та зв'язку: зб. наук. пр., Київ, ЦНДІ навігації і управління, вип. 3(7), с. 85–89.