

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ  
СХІДНОУКРАЇНСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМ. В. ДАЛЯ  
ФАКУЛЬТЕТ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ ТА ЕЛЕКТРОНІКИ  
КАФЕДРА КОМП'ЮТЕРНИХ НАУК ТА ІНЖЕНЕРІЇ

До захисту допускається  
в.о. завідувача кафедри  
\_\_\_\_\_ Рязанцев О.І.  
« \_\_\_\_ » \_\_\_\_\_ 20\_\_ р.

**МАГІСТЕРСЬКА РОБОТА**

НА ТЕМУ:

**Моделі і методи класифікації зображень для систем комп'ютерного зору**

---

---

---

Освітній рівень “Магістр”  
Спеціальність 122 “Комп’ютерні науки”

Науковий керівник роботи:

\_\_\_\_\_  
(підпис)

Д.О.Недзельський  
(ініціали, прізвище)

Консультант з охорони праці:

\_\_\_\_\_  
(підпис)

Я.О.Критська  
(ініціали, прізвище)

Студент:

\_\_\_\_\_  
(підпис)

Я.О.Коробка  
(ініціали, прізвище)

Група:

КН-19дм

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ  
СХІДНОУКРАЇНСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ  
ІМЕНІ ВОЛОДИМИРА ДАЛЯ

Факультет Інформаційних технологій та електроніки  
Кафедра Комп'ютерних наук та інженерії  
Освітній рівень магістр  
Напрямок підготовки \_\_\_\_\_  
(шифр і назва)  
Спеціальність 122 "Комп'ютерні науки"  
(шифр і назва)

**ЗАТВЕРДЖУЮ:**

Т.в.о. завідувача кафедри \_\_\_\_\_  
В.С.Кардашук  
« \_\_\_\_\_ » \_\_\_\_\_ 20 \_\_\_\_ р.

**З А В Д А Н Н Я  
НА МАГІСТЕРСЬКУ РОБОТУ СТУДЕНТУ**

Коробці Ярославу Олександровичу  
(прізвище, ім'я, по батькові)

1. Тема роботи Моделі і методи класифікації зображень для систем комп'ютерного зору

керівник проекту (роботи) Недзельський Дмитро Олександрович, к.т.н., доц.  
(прізвище, м.я, по батькові, науковий ступінь, вчене звання)

затверджені наказом вищого навчального закладу від «5» 10 2020 р. № 140/15.15

2. Строк подання студентом роботи 10.01.2021

3. Вихідні дані до роботи Матеріали науково-дослідної практики, математичні моделі виділення ключових точок, апарат мультимножин, методи кодування із використанням функцій Уолша, перелік використовуваних програмних засобів: мова програмування python, бібліотеки програмного забезпечення

4. Зміст розрахунково-пояснювальної записки (перелік питань, які потрібно розробити) Аналіз методів інваріантного розпізнавання зображень, трансформація простору ознак з метою підвищення ефективності розпізнавання, мультимножини у структурному розпізнаванні, результати експериментальних досліджень, охорона праці та безпека в надзвичайних ситуаціях, висновки

5. Перелік графічного матеріалу (з точним зазначенням обов'язкових креслень)  
Електронні плакати

## 6. Консультанти розділів проекту (роботи)

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
Охорона праці та безпека в надзвичайних ситуаціях	Критська Я.О. ст. викл. кафедри КНІ		

7. Дата видачі завдання 14.10.2020

Керівник

\_\_\_\_\_ (підпис)

Завдання прийняв до виконання

\_\_\_\_\_ (підпис)

## КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів дипломного проекту (роботи)	Строк виконання етапів проекту ( роботи )	Примітка
1	Розробка технічного завдання	02.09.2020-15.09.2020	
2	Аналіз літератури з досліджуваної проблеми	16.09.2020-22.09.2020	
3	Розробка методу та моделі	23.09.2020-25.09.2020	
4	Реалізація методу	26.09.2020-06.10.2020	
5	Аналіз результатів дослідження	07.10.2020-25.11.2020	
6	Розробка частини проекту "Охорона праці та безпеки в надзвичайних ситуаціях"	26.11.2020-1.12.2020	
7	Оформлення пояснювальної записки, автореферату та презентації	2.12.2020-09.01.2021	

Студент

\_\_\_\_\_ (підпис)

Коробка Я.О.

\_\_\_\_\_ (прізвище та ініціали)

Науковий керівник

\_\_\_\_\_ (підпис)

Недзельський Д.О.

\_\_\_\_\_ (прізвище та ініціали)

## АНОТАЦІЯ

Коробка Я.О. Моделі і методи класифікації зображень для систем комп'ютерного зору.

Об'єктом дослідження є методи розпізнавання зображень у системах комп'ютерного зору.

Метою дослідження є розроблення методів зіставлення структурних описів зображень на основі апарату мультимножин та застосування результативних трансформацій описів.

Основним завданням прийнято розробку методів обчислення релевантності описів зображень із використанням апарату мультимножин, проведення експериментальних досліджень розроблених методів, порівняння результатів з відомими методами для конкретних баз візуальних даних. Для отримання бінарних дескрипторів ключових точок застосовано детектор ORB.

У результаті роботи здійснена програмна реалізація методу для обчислення релевантності зображень.

**Ключові слова:** комп'ютерний зір, структурні описи зображення, ключова точка, дескриптор, мультимножини, функції Уолша.

## ABSTRACT

Korobka Ya.O. Models and methods of image classification for computer vision systems.

The object of the study is the methods of image recognition in computer vision systems.

The purpose of the study is to develop methods for comparing structural descriptions of images on the basis of the multisets apparatus and to apply effective transformations of descriptions.

The main task was to develop methods for calculating the relevance of image descriptions using the multisets apparatus, to conduct experimental studies of the developed methods, to compare the results with known methods for specific visual data bases. An ORB detector was used to obtain binary keypoint descriptors.

As a result, a software implementation of the method for calculating the relevance of the images was carried out.

**Keywords:** computer vision, structural descriptions of image, key point, descriptor, multisets, Walsh functions.

## ЗМІСТ

ПЕРЕЛІК СКОРОЧЕНЬ ТА УМОВНИХ ПОЗНАЧЕНЬ .....	5
ВСТУП.....	6
1 АНАЛІЗ МЕТОДІВ ІНВАРІАНТНОГО РОЗПІЗНАВАННЯ ЗОБРАЖЕНЬ.....	8
1.1 Структурні методи розпізнавання.....	8
1.2 Ознакові описи як множина дескрипторів ключових точок .....	13
1.2.1 Технології SIFT та SURF .....	14
1.2.2 Технології FAST та BRIEF, ORB.....	18
1.3 Обчислення релевантності описів.....	21
1.4 Постановка задачі дослідження.....	25
2 ТРАНСФОРМАЦІЯ ПРОСТОРУ ОЗНАК З МЕТОЮ ПІДВИЩЕННЯ ЕФЕКТИВНОСТІ РОЗПІЗНАВАННЯ .....	26
2.1 Стиснення опису на основі застосування ортогонального базису функцій Уолша...26	
2.2 Дослідження властивостей трансформованих описів.....	33
2.3 Компресування опису візуальних об'єктів шляхом фільтрації його компонент за критерієм стабільності .....	37
2.4 Оптимізація параметра порога при зіставленні описів об'єктів.....	43
2.5 Мультимножини у структурному розпізнаванні .....	47
3 РЕЗУЛЬТАТИ ЕКСПЕРИМЕНТАЛЬНИХ ДОСЛІДЖЕНЬ.....	52
3.1 Обґрунтування вибору середовища програмної реалізації .....	52
3.2 Результати експериментальних досліджень.....	53
4 ОХОРОНА ПРАЦІ .....	58
4.1 Аналіз небезпечних та шкідливих факторів при виробництві (експлуатації) виробу58	
4.2 Гігієнічні вимоги до параметрів виробничого середовища.....	60
4.2.1 Мікроклімат .....	60
4.2.2 Освітлення .....	60
4.2.3 Шум та вібрація, електромагнітне випромінювання.....	62
4.2.4 Вентилювання .....	63
4.3 Заходи з організації виробничого середовища та попередження виникнення надзвичайних ситуацій.....	63
ВИСНОВКИ.....	67
ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАНЬ .....	68
ДОДАТОК А. Фрагмент лістингу.....	73
ДОДАТОК Б. Електронні плакати.....	81

## ПЕРЕЛІК СКОРОЧЕНЬ ТА УМОВНИХ ПОЗНАЧЕНЬ

SIFT – Scale-Invariant Feature Transform

SURF – Speeded Up Robust Features

ORB – Oriented FAST and Rotated BRIEF

НС – найближчі сусіди

ІПО – інваріантний признаковий опис

ІРЗ – інваріантне розпізнавання зображень

МХО – множина характерних ознак

ОО – ознакові описи

ПУА – перетворення Уолша-Адамара

ФУ – функції Уолша

ХО – характерна ознака

## ВСТУП

Однією з найактуальніших і ключових для сучасних систем комп'ютерного зору проблем є інваріантне розпізнавання зображень на основі побудови їх ознакових описів (ОО). Її рішення потрібно для прикладних задач, наприклад, таких як визначення просторових координат об'єктів методами зондування, автоматизація стеження за рухомими об'єктами, класифікація і пошук зображень в базах відеоінформації, аналіз і прийняття рішень на основі візуальної інформації в медичних системах, обробка та аналіз зображень при різного роду мікроскопічних, космічних та інших дослідженнях.

Існує різноманіття ознак об'єктів, які можуть мати різну природу і значимість для розпізнавання, тому їх відбір за ступенем впливу на якість результату часто набуває вирішального значення. Швидкість обробки значно підвищує використання описів зображень для зберігання і аналізу, а не самих зображень, за рахунок відсутності етапу первинної обробки. Основною проблемою при цьому часто є надмірність сформованих ОО, і, як наслідок, втрата швидкодії при розпізнаванні. Для описів у вигляді множин структурних ознак це обумовлено тією обставиною, що вони зазвичай містять підмножини точок, які при геометричних перетвореннях не можуть бути поставлені у відповідність ніяким структурним ознаками з аналізованих зображень. Множини структурних ознак, побудовані найбільш популярними в застосунках методами SURF, SIFT для дискретних зображень, наприклад, розміром  $640 \times 480$  можуть містити до 1000 точок-дескрипторів. На практиці надмірність призводить до втрати показника швидкодії розпізнавання і ставить проблему отримання і вивчення неповних описів [1-5].

Вирішенням проблеми є розробка нових ефективних моделей і методів перетворення ОО при аналізі зображень, які дозволяють забезпечити необхідний рівень швидкодії без втрати якості розпізнавання. Застосовування методу зіставлення структурних описів зображення з використанням мультимножин може істотно сприяти підвищенню швидкодії і поліпшенню інших характеристик класифікації за рахунок використання процедур стиснення опису або переходу в простір ознак, що забезпечує найкращі показники розпізнавання.

Також існує проблема побудови ефективних мір подібності структурних описів у вигляді множин на основі селекції в процесі її обчислення найбільш значущих інформативних компонентів опису.

Такі міри можуть бути визначені як шляхом відбору найбільш близьких до еталонних ознак об'єкта на основі принципу найближчих сусідів, так і на основі прийняття

рішення про подібність з використанням кінцевого підмножини найбільш істотних для розпізнавання елементів опису.

Виникає необхідність вирішення завдань адаптивного або оптимального вибору параметрів методів, таких як пороги еквівалентності для елементів опису, у зв'язку з розширенням сфер практичного застосування методів структурного розпізнавання [6-10].



# 1 АНАЛІЗ МЕТОДІВ ІНВАРІАНТНОГО РОЗПІЗНАВАННЯ ЗОБРАЖЕНЬ

## 1.1 Структурні методи розпізнавання

Останнім часом при побудові прикладних систем зорового сприйняття все більшого поширення набувають структурні методи, у яких система ознак для здійснення розпізнавання будується на основі множини характеристик окремих ключових точок зображення – множина характерних ознак (МХО).

Вихідною моделлю простору ознак опису для таких методів є скінченна множина  $Z \subset \mathbb{R}^n$  векторів, кожен з яких пов'язаний з точкою просторових координат  $(x, y) \in \mathbb{C}^2$ . Саме за допомогою структурних методів на основі ОО у вигляді МХО можна вирішити такі практичні завдання, як виділення заданого об'єкта в сцені на складному фоні інших об'єктів, розпізнавання візуального об'єкта по його окремим видимим частинам.

Яскравим прикладом складної сцени і одночасно бази еталонів є набір зображень осіб – ілюстрація, представлена сайтом Google до Міжнародного жіночого дня 8 березня 2019 року (рис.1.1).

Для цієї бази (для кожного окремо взятого зображення) за допомогою методу SURF було отримано опис у вигляді МХО, а потім обчислена матриця подібності між усіма зображеннями бази для різних значень параметрів методу (зокрема, порога для еквівалентності двох векторів опису) без урахування дії геометричних перетворень [2].

Як показав експеримент, кожне з представлених 27-ми зображень бази легко розпізнається серед інших в сцені рисунку 1.1. Число ХО, з яких складаються зформовані описи еталонів, коливається в межах від 14 до 53. Як приклад, число ХО, еквівалентних між першим і іншими зразками, не перевищує 23% (6 точок з 26 для першого зразка).



Рисунок 1.1 – Ілюстрація сайту Google

Картинка (рис. 1.1) є практично ідеальною з точки зору ефективності застосування методу SURF для класифікації об'єктів в обмеженій базі даних. Це обумовлено тим фактом, що множина ознак еталонів тут мають незначний перетин між собою, тобто в базі немає дуже схожих між собою з точки зору застосовуваної міри подібності зображень, що мають в складі своїх структурних описів ідентичні векторні ознаки. Важко собі уявити, що традиційні методи, засновані, наприклад, на інтегральних характеристиках зображення, зможуть впевнено знайти довільний еталон серед усієї множини об'єктів, представлених на зображенні Google. Це може бути можливим лише в ситуації, якщо шуканий об'єкт виділено з фону інших об'єктів.

Теорія інваріантного розпізнавання зображень (ІРЗ) розвивається з 60-х років 20-го століття як теоретичне обґрунтування створюваних прикладних систем комп'ютерного зору.

Інваріантність ознакового опису зображення будемо розуміти наступним чином. Нехай  $G$  – група геометричних перетворень,  $g \in G$  – її елементи,  $B$  – зображення об'єкта з простору  $B \in W$ ,  $gB$  – перетворене зображення. Опис  $Z(B)$  називають інваріантним, якщо виконана умова:

$$\forall g \in G, \forall B \in W : Z(B) = Z(gB), \quad (1.1)$$

де рівність  $Z(b)=Z(gB)$  розуміється в сенсі еквівалентності множин або векторів, тобто ознаки зображень з одного класу еквівалентності збігаються [3]. Саме властивість інваріантності поряд з компактністю опису, здатністю адаптації та автоматичного аналізу є ключовим для сучасних прикладних систем розпізнавання, заснованих на інтелектуальному аналізі даних.

Разом з інваріантністю при використанні описів МХО виконана також важлива з прикладної точки зору вимога перешкодозахищеності, тобто збереження здатності функціонування в умовах впливу деякого рівня перешкод. Відомі два найбільш поширених підходу до ІРЗ:

- 1) побудова системи інваріантних ознак і зіставлення їх з еталонними значеннями;
- 2) нормалізація зображення, тобто обчислення значень параметрів геометричних перетворень і відбудова від них шляхом компенсації.

Обидва напрямки в традиційному варіанті побудовані на ознаках інтегрального виду:

$$\Phi(B, f, K) = \iint_E f[B(x, y)]K(x, y)dx dy, \quad (1.2)$$

де  $E(x, y)$  – двовимірна область визначення функції яскравості зображення  $V(x, y)$ ;

$f$  – функція перетворення зображення,

$K$  – функція ядра інтегрального функціоналу (1.2), що визначає його властивості.

Ознаки  $\Phi(V, f, K)$  мають виключно важливу властивість стійкості до аддитивному шуму, що важливо в ряді технічних програм, однак, їх недоліком (особливо при використанні в нелінійних обчисленнях) вважається чутливість до перешкод дискретизації і локальних перешкод у вигляді фону і помилкових об'єктів [4].

Функціонали (1.2) з успіхом застосовуються також в сучасних методах, пов'язаних з аналізом проєкцій об'єктів. Можна вважати, що функціонали (1.2) з потрібними властивостями використовуються і в сучасних структурних методах, де інтегрування здійснюється в межах локальної області – фрагменті зображення.

Основними проблемами застосування структурних методів ІРЗ є надійне формування вихідного опису у вигляді МХО, що характеризують окремі ділянки зображення, а також побудова вирішального правила для віднесення отриманого опису до ідеалу. Спочатку основою побудови структурних підходів була теорія формальних граматики. Основою будь-якого методу розпізнавання є міра подібності з еталонним описом, від побудови якої безпосередньо залежить визначення класу об'єкта. Найбільшого поширення в просторах різного типу ознак, в тому числі і при вирішенні проблем ІРЗ, отримали метричні, кореляційні, статистичні підходи, ідейна база яких закладена і в принципах побудови сучасних структурних підходів.

Метричні методи пов'язують з використанням метрик у процедурах зіставленні описів, що забезпечує істотний вигравш у часі в процедурах пошуку в базах зображень.

Кореляційні принципи розпізнавання зображень залишаються затребуваними завдяки своїй винятковій прикладної працездатності в складних умовах і знайшли своє застосування в сучасних прізнакових методах аналізу відеоінформації. Стосовно до розпізнавання в умовах дії групи перетворень  $G$  суть кореляційного методу зводиться до наступного: Нехай  $V = gV_0$  – вхідне зображення,  $g \in G$  – елемент групи,  $V_0$  – еталон. Кореляційний підхід зводиться до послідовного перебору всіх значень параметрів групи  $G$  з метою аналізу коефіцієнта кореляції

$$\xi(V, V_0) = \frac{(V, V_0)}{\|V\| \cdot \|V_0\|}, \quad (1.3)$$

де  $(\mathbf{B}, \mathbf{B}_0)$  – скалярний добуток зображень як векторів, яке можна обчислити як

$$(\mathbf{B}, \mathbf{B}_0) = \sum_{i=1}^N \sum_{j=1}^N B(i, j) B_0(i, j);$$

$\|\mathbf{B}\|, \|\mathbf{B}_0\|$  – норми  $\mathbf{B}, \mathbf{B}_0$ .

Рішення приймається по максимальному або близьким до одиниці значенням коефіцієнта  $\xi(\mathbf{B}, \mathbf{B}_0)$ , який приймає значення  $|\xi(\mathbf{B}, \mathbf{B}_0)| \leq 1$  [5].

Використання кореляції засноване на переборі всіх відповідних вхідному зображенню еталонів, з огляду на їх можливі перетворення. Одним із суттєвих недоліків даного методу є недостатня швидкодія, яке іноді виходить за рамки допустимого при реалізації в режимі реального часу. Іноді цей недолік можна згладити за допомогою зберігання необхідних даних в пам'яті пристроїв розпізнання.

Підхід до IPЗ за допомогою ще одного поширеного підходу – штучних нейронних мереж функціонально заснований на здатності нейронів живих організмів самонавчатися, що знайшло застосування в задачах комп'ютерного зору. Нейронні мережі, що складаються з наборів шарів нейронів, можуть бути навчені складній структурі зразків в багатовимірному просторі. Крім того, нейронні мережі допускають реалізацію в формі паралельних алгоритмів [6].

Штучний нейрон є акумулятором вхідних сигналів з відповідними ваговими коефіцієнтами. Вихідний сигнал нейрона надсилається за єдиним аксоном, який розгалужується на множину сигналів для входу на наступні шари мережі. Вихідний сигнал формується застосуванням до зваженої суми деякої функції.

Навчання нейрона виробляється за допомогою підстроювання ваг сигналів, що подаються на його вхід для отримання відомого вихідного сигналу. Розроблено стандартні алгоритми навчання нейронної мережі. Обчислювальні можливості нейрона є великий теоретичний і практичний інтерес. Відомо, що штучний нейрон може моделювати логічні функції І, АБО і НЕ. У той же час, будь-яку логічну функцію можна реалізувати шляхом об'єднання кількох штучних нейронів. За допомогою апарату нейронних мереж вирішуються такі актуальні інтелектуальні завдання, як навчання і моделювання, розпізнавання образів, апроксимація функцій і багато інших. Відомі застосування нейронних мереж для вирішення задач верифікації підписів, розпізнавання осіб і пошуку пішоходів.

Найбільш принциповими недоліками використання нейронних мереж при геометричних перетвореннях зображень є необхідність реалізації об'ємного процесу навчання, який може займати тривалий час через велике розмаїття можливих ситуацій. Крім того, ефективність застосування методу в значній мірі залежить не тільки від

структури мережі, а й від кількості нейронів в кожному шарі і числа шарів, які можуть варіюватися в залежності від завдання.

Група статистичних методів безпосередньо пов'язана з розглянутими підходами. Оптимізована в базі еталонів міра подібності описів може бути представлена як значення статистичного критерію, де вектори описів виступають в ролі вимірювань. Застосування оптимальної статистичної теорії розпізнавання найбільш ефективно в разі відомих імовірнісних описів класів. У більшості завдань комп'ютерного зору імовірнісні розподіли значень ознак важко уявити аналітично, тому класичне застосування статистичних критеріїв ускладнено, хоча елементи такої обробки використовуються в сучасних методах SIFT, SURF. З цих причин основним критерієм при розпізнаванні можна вважати значення міри подібності описів [7].

У найзагальнішому вигляді класифікація об'єкта на основі опису у вигляді МХО  $Z$  є відображення  $\mathfrak{R}: Z \rightarrow \{Z(j)\}$ , где  $\{Z(j)\}_{j=1}^J$  – кінцева множина еталонних описів (база відеоданих). Відображення  $\mathfrak{R}$ , як правило, реалізується шляхом вирішення задачі оптимізації:

$$c(Z) = \arg \operatorname{opt}_{j \in \{1, \dots, J\}} \mathfrak{G}[Z, Z(j)], \quad (1.4)$$

де  $\mathfrak{G} \in \mathbf{R}^1$  – міра подібності;

$c(Z)$  – функція отримання класу об'єкту.

Відомі методи зіставлення ОО засновані на процедурі голосування, вони зосереджуються на обчисленні та оптимізації по параметру  $j$  міри  $v[Z, Z(j)]$  у вигляді значення частки елементів опису  $Z$ , знайшли відповідність (еквівалентність) в  $j$ -му еталоні.

Двома головними факторами, від яких вирішальною мірою залежить якість класифікації, як відомо, є подібність елементів всередині класів і відмінності між елементами різних класів. У статистичній теорії класифікації ці фактори позначаються значеннями критеріїв всередині класового розкиду і міжкласових відстаней.

Одним з перспективних і обговорюваних в літературі підходів до класифікації на основі групового прийняття рішення про клас або про подібність на основі множини часткових рішень є бустінг (boosting), зокрема, це застосовується задачі розпізнавання осіб. Тут використовується близький до нейронних мереж принцип лінійної комбінації локальних рішень [8].

Ідея методу полягає в побудові серії «простих» класифікаторів, рішення яких через обмежений рівень достовірності не можуть використовуватися як остаточний варіант, але на їх основі має сенс побудова інтегрованого класифікатора у вигляді зваженої суми результатів часткових рішень. Агрегований бінарний класифікатор  $h(x)$  для випадку двох класів може бути побудований як зважена сума простих класифікаторів  $h_j$

$$h(x) = \text{sign} \left[ \sum_{j=0}^{m-1} \alpha_j h_j(x) \right], \quad (1.5)$$

де  $\alpha_j$  – вага простого класифікатора  $h_j(x)$ ;

$m$  – кількість часткових класифікаторів.

У варіантах boosting у ролі простих класифікаторів виступають порогові функції

$$h_j(x) = a_j [f_j < \theta_j] + b_j [f_j \geq \theta_j] = \begin{cases} a_j, & \text{if } f_j < \theta_j, \\ b_j, & \text{otherwise,} \end{cases} \quad (1.6)$$

які також відомі як гілки рішень (дерева рішень).

У більшості випадків встановлюють величину  $a_j, b_j$  рівними  $\pm 1$ , тому необхідно визначити тільки функцію  $f_j$  і межове значення  $\theta_j$ .

Розглядаючи рішення про клас окремих ХО із структурного опису у вигляді МХО як результат застосування часткових класифікаторів, агреговане рішення  $h(x)$  можна представити у вигляді композиції часткових рішень [9].

## 1.2 Ознакові описи як множина дескрипторів ключових точок

На рисунку 1.2 наведено приклад перетвореного зображення із досліджуваної бази риб, поданого у вигляді точок з координатами дескрипторів ОО з МХО.

Найчастіше застосовуються інваріантно-ознакові методи SIFT, SURF. Вони дозволяють сформувати і зіставити набір ХО на вхідному і еталонному зображеннях, при цьому вхідне зображення може бути піддано впливу спотворень підгруп афінної групи.



Рисунок 1.2 – Приклад перетвореного зображення з координатами точок МХО

### 1.2.1 Технології SIFT та SURF

Технологія SIFT (Scale-Invariant Feature Transform – масштабно-інваріантне перетворення ознак) є інструментом обчислення ППО зображень. Головною особливістю ОО в методі SIFT є стійкість (інваріантність) до геометричних перетворень.

В процесі багатоетапної обробки методом SIFT шляхом посилення відмінностей в просторі ознак досягається необхідна величина відстані між структурними елементами різних об'єктів, що в цілому призводить до високих характеристик розпізнавання [10].

Значення ХО зображення відображають особливості фрагментів. Основним застосуванням методу є розпізнавання двовимірних об'єктів (включаючи частково закритих) і реконструкції тривимірних сцен.

Початкове зображення перетворюється до виду множини функцій:

$$L(x, y, s) = B(x, y) \otimes W(x, y, s), \quad (1.7)$$

де  $B(x, y)$  – функція яскравості зображення;

$W(x, y, s)$  – Гауссіан з параметром  $s$ ;

$\otimes$  – символ згортки.

Далі на базі системи функцій  $L$  формується множина:

$$D(x, y, s_i) = L(x, y, s_{i+1}) - L(x, y, s_i). \quad (1.8)$$

Вираз (1.8) являє собою апроксимацію оператора Лапласа. Локальні екстремуми функції (1.8) формують множина ХО.

Для уточнення положення кожної точки в просторі координат здійснюють розкладання  $D(x, y, S_i)$  в ряд Тейлора в характерних точках:

$$D(c) = D + \frac{\partial D^T}{\partial c} c + \frac{1}{2} c^T \frac{\partial^2 D}{\partial c^2} c, \quad (1.9)$$

де  $c = (x, y)$  – зміщення від обраних ХО  $z^*$  в межах фіксованого масштабу  $S_i$ . Місцезнаходження  $\hat{c}$  екстремуму функції  $D(c)$  по зсуву визначається шляхом прирівнювання нулю похідної від  $D(c)$  щодо  $c$ . Після цього положення детектора ще уточнюється.

Формування ХО в точці здійснюється шляхом дослідження співвідношення між головними значеннями кривизни в напрямках  $x, y$ . Визначення параметрів кривизни

зводиться до знаходження власних значень матриці гессіан  $H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix}$ , де  $D_{uv}$  –

друга похідна функції  $D$  по змінним  $u, v$ . Власні значення  $\alpha_1, \alpha_2$  матриці  $H$  пропорційні просторової кривизни функції  $D$  при фіксованому  $S_i$ . Реалізація цього етапу виконано за допомогою рішення системи рівнянь:

$$\alpha_1 \alpha_2 = \text{Det}(H), \quad \alpha_1 + \alpha_2 = \text{Tr}(H), \quad (1.11)$$

де  $\text{Tr}$  – слід, а  $\text{Det}$  – визначник матриці.

Оціночна функція має вигляд:

$$R = \det(H) - k \text{Tr}(H)^2, \quad (1.12)$$

де  $k$  – емпірична константа,  $k = 0.04 - 0.06$ . Крайовими точками вважають ті, для яких  $R < 0$ .

Кожній точці призначається одна або кілька орієнтацій на основі локальних значень градієнта [11]. Для досягнення інваріантності до обертань щодо цієї орієнтації в подальшому будуються дескриптори точок.



Для зображення  $L(x, y, s_i)$  при фіксованому значенні  $s_i$  обчислюються амплітуда  $m(x, y)$  і орієнтація  $\theta(x, y)$  градієнта для ХО з координатами  $(x, y)$

$$m(x, y) = \sqrt{[L(x+1, y) - L(x-1, y)]^2 + [L(x, y+1) - L(x, y-1)]^2}, \quad (1.13)$$

$$\theta(x, y) = \arctan \frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)}. \quad (1.14)$$

Ознакова інформація міститься в масиві напрямків  $T = \{t_i\}$ ,  $t_i = (x, y, m, \theta, s_i)$ , обчисленому для МХО. В околиці  $16 \times 16$  кожної точки при фіксованому масштабі аналізу набору з 16 непересічних фрагментів  $4 \times 4$ , утворюють розбиття. Для кожного з фрагментів, в свою чергу, будується 8-містовцова гістограма напрямків по змінній  $\theta$ , охоплює 360 градусів.

Один напрямок відповідає діапазону в 40 градусів. Значеннями дескриптора є суми амплітуд  $v_i = \sum_{q=1} m_q$ . В результаті для кожної характерної точки маємо вектор (дескриптор), що містить  $16 \times 8 = 128$  значень напрямків.

Для забезпечення інваріантності до поворотів напрямки нормалізуються шляхом зсуву щодо направлення ХО, а для забезпечення стійкості до зміни освітлення гістограми нормалізуються шляхом ділення компонент на норму вектора [15].

Більш сучасний метод SURF вирішує аналогічні два завдання – пошук особливих точок зображення і створення їх дескрипторів, інваріантних до масштабу і обертання. Градієнт в точці в ньому обчислюється за допомогою фільтрів Хаара. Дескриптор SURF (значення ХО) являє собою набір з 64 (або 128) чисел для кожної ключової точки. Ці числа відображають флуктуації градієнта навколо ключової точки. Оскільки ключова точка являє собою максимум функції гесіан, то це гарантує, що в околиці точки повинні бути ділянки з різними значеннями градієнта.

Для інваріантності обчислення дескрипторів потрібно визначити переважну орієнтацію перепадів яскравості. Це поняття близьке до поняття градієнта, але SURF в порівнянні з SIFT використовує інший алгоритм знаходження вектора орієнтації. Обчислюються точкові градієнти в пікселях, сусідніх з особливою точкою. Для розгляду беруться пікселі в окружності радіуса  $6s$  навколо особливої точки, де  $s$  – масштаб. Для обчислення градієнта використовується фільтр Хаара розміром  $4s$ . Фільтри Хаара дають точкове значення перепаду яскравості по осях координат.

Дескриптор вважається як опис градієнта для 16 квадрантів навколо особливої точки. Характерне для SURF зниження ймовірності розпізнавання при збільшенні кута огляду на 45 градусів пов'язано з тим, що SURF НЕ інваріантний до афінних перетворень. Стійкий результат спостерігається при обертанні, зсувах і зміні масштабу зображення.

SURF знайшов застосування для пошуку об'єктів на зображенні, однак він ніяк не виділяє об'єкт з фону. SURF розглядає зображення як єдине ціле і шукає його особливості. При цьому особливості можуть бути як всередині об'єкта, так і на тлі, а також на точках межі об'єкту і фону. У зв'язку з цим метод буває не зовсім ефективним для об'єктів простої форми і без яскраво вираженої текстури.

Усередині таких об'єктів визначається занадто мало особливих точок, точки часто знаходяться або на кордоні об'єкта з фоном, або тільки на тлі. А це може призвести до того, що об'єкт не зможе бути розпізнаний [16].

Ще один метод – ASIFT містить в якості основи технологію SIFT, але додатково забезпечує інваріантність до зміни кута, що визначає осі орієнтації камери, що знімає об'єкт. Метод дає збіги на парі зображень в афінних системах координат, де SIFT не працює.

До недоліків ASIFT відносять невисоку стійкість при обробці великорозмірних зображень, де SIFT стійкий, а також низьку продуктивність через зростання обсягу обчислень при моделюванні афінних спотворень.

В цілому можна відзначити, що методи SIFT, SURF, ASIFT дають надійні результати при вирішенні широкого кола прикладних задач комп'ютерного зору. Основний ідеологією побудови цих методів можна вважати часто застосовуються на практиці кореляційний і статистичний підходи. В цілому SURF вважають сучасною апроксимацією і удосконаленням раніше виниклого методу SIFT. Стандартна версія методу SURF в кілька разів швидше, ніж SIFT, і більш стійка до трансформацій зображення.

## 1.2.2 Технології FAST та BRIEF, ORB

FAST алгоритм працює наступним чином. Піксель  $p$  у масиві порівнює яскравість  $p$  з оточуючими 16 пікселями, які знаходяться в малому колі навколо  $p$ . Пікселі в колі потім сортують у три класи (світліші за  $p$ , темніші за  $p$  або подібні до  $p$ ) [17].

Якщо більше 8 пікселів темніше або яскравіше  $p$ , то його обрано як ключову точку. Таким чином, ключові точки знайдені швидко, дають нам інформацію про розташування визначальних ребер на зображенні [17].

BRIEF описує всі ключові точки, знайдені за допомогою FAST алгоритму, і перетворює їх у бінарні векторні ознаки, щоб вони разом могли представляти об'єкт. Вектор двійкових функцій також відомий як двійковий дескриптор ознак – це бінарний вектор, який містить лише 1 і 0. Кожна ключова точка описується вектором, що становить 128–512 біт рядка.

BRIEF починає зі згладжування зображення за допомогою ядра Гаусса, щоб запобігти чутливості дескриптора до високочастотного шуму. Тоді BRIEF вибирає випадкову пару пікселів у визначеному сусідстві навколо цієї ключової точки. Визначене сусідство навколо пікселя відоме як патч, який є квадратом певної ширини та висоти пікселів. Перший піксель у випадковій парі виводиться з Гауссового розподілу, зосередженого навколо ключової точки, з нитковим відхиленням або поширенням сигми.

Другий піксель у випадковій парі виводиться з розподілу Гаусса, зосередженого навколо першого пікселя, зі стандартним відхиленням або розворотом сигми на два. Тепер, якщо перший піксель яскравіший за другий, він присвоює значення 1 відповідному біту [17].

ORB – це дуже швидкий бінарний дескриптор на основі BRIEF, який є інваріантним обертанню і стійким до шуму. Це можна продемонструвати за допомогою експериментів, в яких ORB на порядок швидше, ніж SIFT, в той же час успішніший в багатьох ситуаціях.

ORB спирається на добре відомий ключовий пункт FAST детектор і нещодавно розроблений BRIEF дескриптор; з цієї причини ми називаємо цей метод ORB (Oriented FAST та Rotated BRIEF). Обидва ці методи є привабливі завдяки їх хорошим експлуатаційним характеристикам та низькому рівню вартості. ORB включає швидкість і точність, ефективне обчислення орієнтованого BRIEF, аналіз дисперсії та співвідношення орієнтованих ознак BRIEF.

ORB також використовує метод навчання для де-кореляції BRIEF особливості при обертальній інваріантності, що призводить до кращої продуктивності у програмах найближчих сусідів.

Дослідження дипломної роботи базуються в основному на застосуванні методу ORB [20].

Сучасним розвитком і узагальненням уявлення системи ознак у вигляді МХО є поняття «мішок ознак» (bag-of-features).

Ці методи продемонстрували високий рівень продуктивності в рішенні задач пошуку в індексованих базах даних зображень.

Суть таких методів полягає в тому, що замість зберігання набору 128-мірних дескрипторів для кожного зображення зберігання підлягає тільки один запис у вигляді «візуального слова», яке описує індексовані зображення. Крім того, використання інвертованих файлів дає можливість порівнювати вектора, беручи до уваги тільки їх ненульові значення, що ефективно в разі значної розрідженості векторів.

Основним недоліком цих підходів є те, що вони не використовують інформацію про просторове розміщення ХО. До речі, традиційний метод SURF також безпосередньо не спирається на координатну інформацію при зіставленні описів, хоча при визначенні координат ХО ґрунтується на фрагменті зображення [21].

Основними перевагами подання ознак у вигляді «мішка» є:

- 1) компактність опису, і, отже, зниження обчислювальних вимог до зберігання та обробки;
- 2) швидкість пошуку в зв'язку з використанням стиснення по методу інвертованих файлів.

В якості альтернативи «мішка ознак» вивчається підхід, який базується на пірамідальному порівнянні векторів (який, проте, також не використовує координатні просторові співвідношення) з попередньої кластеризацією дескрипторів на 4, 8, 16 рівних частин.

Для кожного рівня розбиття будується гістограма частоти значень ознак, що потрапили у відповідну частину розбиття. Рішення приймається за ознаками, які були включені в один кластер. Описаний підхід комбінує два напрямки аналізу дескрипторів МХО: для побудови, порівняння та прийняття рішень використовуються як частини зображення (з ознаками всередині і збереженням структури), так і описові характеристики всього зображення.

В цілому, незаперечними перевагами обговорюваних в роботі методів на основі ОО у вигляді МХО, які обумовлюють їх широке практичне застосування, можна вважати:

- інваріантність при геометричних перетвореннях повороту, зсувів і масштабу аналізованого об'єкта;
- можливість прийняття рішення в частині об'єкта, що недоступно в підходах інтегрального типу;
- стійкість до впливу складного фону і просторових перешкод;
- досить висока швидкодія розпізнавання;
- порівняно високі загальні показники функціонування (швидкодія-перешкодозахищеність);

У той же час класичний метод SURF і близькі до нього підходи не позбавлені наступних важливих недоліків, що вимагають їх розвитку та удосконалення:

- формування ХО за межами аналізованого візуального об'єкта, що ускладнює аналіз;
- значні коливання числа компонентів в описах різних еталонів, що вимагає застосування нормування;
- значні обсяги ОО, що досягають 200-300 і навіть 1000 ХО для окремих зображень еталонів, що безпосередньо впливає на час пошуку та розпізнавання;
- часто спостерігається для різних баз надмірність МХО, той же рівень якості розпізнавання може бути досягнуто і для значно меншого обсягу ознак;
- повторення значень окремих ознак для різних еталонів, що безпосередньо позначається на достовірності розпізнавання;
- залежність результатів розпізнавання від ряду параметрів: порога на еквівалентність ХО при їх зіставленні, співвідношення глобального та локального максимумів подібності ХО при ухваленні рішення про відповідність ХО, що вимагає розробки інтелектуальних процедур обчислення порогів;
- поява значного числа помилкових ХО при дії геометричних перетворень через вплив помилок дискретизації, це викликає необхідність дослідження інваріантних властивостей ОО в практиці роботи з конкретними відео-базами даних.

Всі названі фактори викликають необхідність удосконалення і подальшого розвитку сучасних структурних підходів з описами у вигляді МХО [22].

### 1.3 Обчислення релевантності описів

Розпізнавання (класифікація) – це віднесення вимірних даних до певного класу. Важливим є розвиток такої системи ознак із загальної маси даних, кількість і якість яких необхідна і достатня для здійснення ефективного вирішення завдання в наявних умовах. Актуальним в прикладному аспекті є проведення досліджень, спрямованих на формування ефективних систем ознак на базі описів, отриманих на основі методів виявлення точкових особливостей зображень, які формують ОО об'єкта у вигляді кінцевої МХО векторів-дескрипторів.

Елементи опису інваріантні до зсувів, масштабів і обертання. Нумерація елементів в МХО  $Z$  представляє ОО та виходить безпосередньо в процесі формування опису [24].

Аналіз наукової літератури показує, що в останній час основними шляхами підвищення ефективності розпізнавання на основі наявної системи ознак є або трансформація цієї системи в набір ефективніших в деякому плані ознак, або фільтрація (просіювання) ОО з метою відбору більш концентрованого стисненого множини ознак приблизно з тими ж властивостями, що і повний набір.

Особливістю метричних методів є зіставлення описів об'єктів шляхом обчислення значення метрики. Якщо ОО представлено у вигляді МХО, наприклад, в структурному описі об'єктів, то ключовим моментом стає пошук близьких елементів двох множин та формування на цій основі групового рішення у вигляді узагальненої міри подібності.

Вибір міри подібності (або метрики) є принциповим для якості рішення задачі розпізнавання / класифікації в цілому. Проблема побудови міри подібності, найбільш ефективною для конкретних прикладних задач, невпинно перебуває в центрі уваги дослідників. Традиційні підходи, як правило, побудовані на виборі найкращої відповідності серед елементів порівнюваних множин, аналізованих за принципом «кожен з кожним».

Цей принцип лежить, наприклад, в основі методів ближнього і далекого сусідів, метрики Хаусдорфа. Підрахунок числа відповідностей дозволяє побудувати досить просту практичну процедуру голосування для прийняття рішення про клас об'єкту [26].

Методи з застосуванням аналізу описів у вигляді МХО мають досить високу стійкість до аномальних викидів, в тому числі і в просторі використовуваних ознак. Однак у складних ситуаціях істотного пошкодження даних при комплексній дії перешкод адитивного і аномального видів виникає необхідність посилення і розвитку підходів голосування (стосовно МХО) шляхом синтезу методів, в основі яких лежить більш

надійне і універсальне рішення на базі аналізу відібраних якимось чином кортежів групових відповідностей.

Аналіз підмножин відповідностей застосовують на етапі побудови нових мір подібностей у вигляді модифікацій метрики Хаусдорфа, що, на думку авторів, прибирає неоднозначність і покращує перешкодозахищеність при зіставленні кордонів відео-об'єктів. Подоба описів у вигляді числових ознак тут обчислюють на основі обробки і аналізу рейтингу відповідностей. Доцільним є розвиток цієї ідеї в аспекті розробки і дослідження принципу аналізу кінцевих підмножини найкращих відповідностей з МХО як елементів об'єкта й еталона при обчисленні міри подібності, що, безпосередньо впливає на підвищення достовірності та перешкодозахищеності застосування метричної класифікації в цілому.

Ключовими принципами стійкої до викидів інтелектуальної обробки даних є усереднення і ранжування. Якщо використовувати як основу обробки статистичні методи типу LST-оцінок, де одне з рішень базується на ранжированій усіченій вибірці локальних подіб, причому параметром усічення виступає завжди апріорна величина допустимої частки відібраних даних щодо вихідного їх числа. Значення цього параметра на практиці визначається відповідно до граничних рівнем ушкодження опису. Ранжування необхідно для автоматичного виділення, угруповання і відбору потрібних елементів [27].

Нехай  $A, B$  – кінцеві множини з елементами  $A = \{a_i\}_{i=1}^{n_a}, B = \{b_k\}_{k=1}^{n_b}$ ,  $n_a, n_b$  – потужності множин,  $\rho(a_i, b_k)$  – деяка метрика для елементів множин.

У такому випадку відстань Хаусдорфа  $\rho_H(A, B)$  між множинами визначається як:

$$\rho_H(A, B) = \max[\rho(A, B), \rho(B, A)], \quad (1.15)$$

де  $\rho(A, B) = \max_{a \in A} \rho(a, B)$ ,  $\rho(a, B) = \min_{b \in B} \rho(a, b)$ .

У запропонованій модифікації відстані (1.15) на основі LST-обробки значення метрики в правій частині (1.15) обчислюються як лінійна комбінація (усереднення):

$$\rho_s(A, B) = \frac{1}{s} \sum_{i=1}^s \rho(a_i, B), \quad (1.16)$$

фіксованої кількості  $s < n_a$  упорядкованої за зростанням значень відстаней  $\rho(a_i, B)$ , включених до сформованого кортежу  $\rho(a_1, B) \leq \rho(a_2, B) \leq \dots \leq \rho(a_{n_a}, B)$ . Оброблення (1.16) фактично означає відкидання «великих» відстаней, які відносяться до

перешкод, а значення  $S$  визначається допустимим рівнем спотворень. У той же час усереднення в (1.8) знижує і вплив адитивного шуму [17].

Інший спосіб базується на  $M$ -оцінюванні. Для цього вводиться позитивна величина  $\tau$  – поріг для усунення викидів. Перевищені порогом значення метрики  $\rho(a_i, B)$  замінюються на його значення  $\tau$ , інші використовуються в процедурі усереднення без змін. Тут параметр усереднення  $S$  вже не фіксується і змінюється в залежності від вимірювання. Така обробка також сприяє зниженню дії шуму. Відзначимо, однак, що варіанти обробки (1.15), (1.16) призводять до невиконання метричних властивостей, в результаті модифікації отримані міри вже не будуть метриками.

Паралельно розвиваються оптимальні методи зіставлення МХО для моделі «множина-множина», в основі яких лежить оптимізація критерію, як правило, пов'язаного з обчисленням суми відстаней для однозначних відповідностей елементів множин. Оптимальні методи відрізняються більш високими показниками достовірності, проте, в свою чергу, вимагають значних обчислювальних витрат в порівнянні з більш простими методами голосування або обчисленням значень метрики для множин [25].

Аналіз трансформацій систем ознак показує необхідність дослідження і розвитку підходів, побудованих на цілеспрямованій побудові і обробці підмножини найкращих відповідників елементів (як правило, ранжированих деяким чином), що в цілому має сприяти підвищенню достовірності та перешкодозахищенності застосування класифікації за рахунок використання групового аналізу компонентів МХО. Обчислювальні витрати для синтезованих методів за рахунок скорочення ОО будуть істотно менше, ніж для оптимальних підходів, в той же час достовірність класифікації на практиці для них буде вищою, ніж для методів, заснованих на одиночних відповідностях. Принцип формування та аналізу кортежу найбільш значущих відповідностей близький до підходу найближчих сусідів. Відмінною особливістю введення такого оброблення є використання в побудованій мірі істотно скороченого обсягу елементів еталона, що забезпечує стійкість до спотворень опису.

З іншого боку, представляє прикладний інтерес дослідження впливу перешкод різного типу при застосуванні розроблених модифікованих методів для вирішення значно складніших завдань інваріантної класифікації в базах відеоданих.

Важливу роль в розпізнаванні зображень грає властивість інваріантності до геометричних перетворень, що визначається обраним способом формування елементів ОО. Основною вимогою до МХО є забезпечення достатнього рівня стабільності значень обчислених інваріантних ознак при зміні параметрів геометричних перетворень в деяких межах. Як показують дослідження методів розпізнавання, при дії геометричних перетворень спостерігається процес зміни обсягу ОО як в бік скорочення, так і в бік



розширення, викликаний, перш за все, помилками дискретизації. Відзначимо, що при перетвореннях зміщення на дискретну величину такі зміни відсутні [31].

У той же час наші експерименти показують, що досить значна кількість точок ОО стабільно зберігає значення при перетвореннях, що дозволяє впевнено здійснювати розпізнавання навіть при дії перешкод. Виникає необхідність оцінити мінімально допустиме на практиці число точок ПО і розробити процедуру формування стійкої ОО як підмножини повного опису. В цілому застосування будь-яких процедур компресування ПО направлено на зменшення обсягу обчислень і зниження вимог до методів розпізнавання з точки зору швидкодії. Досягнутий вигравш у часі при цьому пропорційний відношенню потужностей старої і нової (стислій) систем ознак.

Треба мати на увазі, що зменшення обсягу шляхом застосування процедур компресування ОО може привести до зниження ймовірності розпізнавання в умовах перешкод.

У процесі дослідження властивостей детектора SURF з'ясувалося, що якість класифікації істотно залежить від порогового значення  $\delta_z$  для метрики  $\rho(a_i, b_k)$ , визначальної еквівалентність двох елементів ОО, представлених у вигляді множини векторів.

В існуючих підходах вибір порогу здійснюється на основі експерименту. Якщо відстань між елементами  $z_1, z_2$  менше порогу, тобто  $\rho(z_1, z_2) \leq \delta_z$ , то  $z_1, z_2$ , вони вважаються еквівалентними між собою.

На основі значення  $\delta_z$  визначається величина подібності між описами як число голосів, відданих елементами опису еталона за розпізнавання об'єкту.

Наприклад, для конкретної бази риб допустимий діапазон порогових значень склав відрізок  $[0,08; 0,8]$ . при порозі  $\delta_z = 0,8$ .

У процесі порівняння описів визначаються всі відповідності точок перетвореного зображення ідеалу, що характеризує виконання властивості інваріантності в повному обсязі, тобто для всіх точок опису. При значенні порогу  $\delta_z < 0,8$  лише для частини точок опису дотримується ця властивість. При цьому для  $\delta_z = 0,08$  число зберігають інваріантність точок не перевищує 5% від загального обсягу ОО еталона, що може викликати проблеми розрізнення об'єктів між собою [34].

## 1.4 Постановка задачі дослідження

Метою дослідження є вивчення методів побудови, аналізу та зіставлення структурних описів зображень на основі апарату мультимножин та застосування результативних трансформацій описів.

Для реалізації мети у магістерській роботі сформульовані і вирішені завдання, які передбачають:

- зниження надмірності МХО шляхом створення неповного або стисненого опису об'єкта, і як результат, скорочення обсягу структурного опису без втрати якості розпізнавання;

- побудову нових покращених мір подібності на основі найбільш значущих елементів опису, що підвищує швидкодію розпізнавання без зниження перешкодозахищеності;

- підвищення ефективності розпізнавання шляхом застосування для зіставлення МХО апарату мультимножин, що інтегрує інформацію про описи;

- результати дослідження використано для експериментального застосування МХО, отриманих застосуванням детектору ORB.

## **2 ТРАНСФОРМАЦІЯ ПРОСТОРУ ОЗНАК З МЕТОЮ ПІДВИЩЕННЯ ЕФЕКТИВНОСТІ РОЗПІЗНАВАННЯ**

З точки зору вимог завдань розпізнавання уявлення об'єкта у вигляді структурного ОО має суперечливий характер. З одного боку, його елементи повинні бути однорідними (вектори з одного відео-об'єкта). З іншого боку, дескриптори описів різних об'єктів в реальних ситуаціях можуть бути еквівалентними або близькими між собою, в результаті виникають труднощі їх класифікації. Крім того, як показують дослідження, в прикладних задачах класифікатором часто досить використовується обмежене (неповне) число елементів опису для успішного прийняття рішення з тим же рівнем помилки.

Ці об'єктивні обставини породжують можливість постановки та вирішення задачі зменшення надмірності шляхом досягнення компактності описів за рахунок їх цілеспрямованих перетворень для більш ефективного зберігання, аналізу і обробки інформації. Ефективне з точки зору деякого критерію уявлення в просторі векторів може бути пов'язано з перетворенням до ортогональному простору. Практичний інтерес через досить просту програмну або апаратну реалізацію викликає застосування в цих цілях ортогонального простору функцій Уолша (ФУ) і перетворень, пов'язаних з ними [35].

### **2.1 Стиснення опису на основі застосування ортогонального базису функцій Уолша**

Опис у вигляді БХО з точки зору вимог до класифікації має в певному сенсі суперечливий характер. Воно є в деякому плані однорідним. З іншого боку, виникають труднощі через те, що вектора з ОО різних об'єктів можуть бути еквівалентними.

Розглянемо напрямок удосконалення системи ознак, пов'язаний з трансформацією вихідного ОО.

Одна з дієвих можливостей пов'язана із застосуванням ортогональних перетворень. Якщо розглядати ОО у вигляді множин числових векторів, то можлива інтерпретація з теорії обробки сигналів, де енергія сигналу визначається як квадрат довжини (або норми) вектора. Для векторного простору також введено поняття скалярного множення, відстані між векторами, а також поняття ортогональності векторів і базису векторного простору, в

якому будь-який вектор може бути єдиним способом представлений у вигляді лінійної комбінації базисних векторів [38].

Якщо базисні вектори мають одиничну довжину, то говорять про ортонормальні бази. У скінченномірному векторному просторі може існувати довільне число ортонормальних базисів.

Зазвичай для представлення та аналізу вибирають базис, що володіє потрібними для вирішення конкретного завдання властивостями. У теорії аналізу зображень поширення набули бази Робертса, Фрея-Чена, Фур'є, Уолша-Адамара.

Застосування двовимірних перетворень в обробці зображень пов'язані з підкресленням окремих властивостей, кодуванням зображень шляхом відкидання незначних за величиною коефіцієнтів перетворення, скороченням розмірності опису для підвищення швидкодії.

Цікавим з прикладної точки зору через досить просту програмну або апаратну реалізацію є використання в якості ортонормального базису сімейства функцій Уолша-Адамара і пов'язаних з ними перетворень.

Дискретні функції Уолша (ФУ), являють собою цілочисельні вектори скінченної розмірності, що складаються з 1 і -1. У практичних застосуваннях досліджуються ФУ з бінарною розмірністю, тому природним виглядає застосування ФУ, наприклад, для векторів SURF, які мають відповідний розмір. Повний набір ФУ утворює ортогональну матрицю Адамара. Наприклад, матриця Адамара розмірністю 4 має вигляд:

$$A = \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}, \quad (2.1)$$

де коефіцієнт  $1/2$  служить для забезпечення властивості ортонормальності. Ортонормальність при практичному застосуванні ФУ грає не таку принципову роль, як, наприклад, при використанні базису Фрея-Чена, тому що нормувальні коефіцієнти рівні для всіх ФУ одного базису. На основі спектрального уявлення можна досягти скорочення опису шляхом формування та аналізу найбільш інформативної частини спектру, за рахунок чого досягається загальне скорочення обчислювальних витрат на обробку в процесі розпізнавання [39].

Мета досліджень – розробка ефективного в плані швидкодії методу класифікації зображень на основі побудови стислих описів з використанням найбільш інформативних компонентів спектра Уолша. Вибір найбільш інформативного кортежу з множини ФУ

здійснюється з використанням дисперсійного критерію, що відображає енергетичні властивості відгуків в обраному базисі.

Для компактності подамо опис у вигляді матриці, рядки якої містять вектори опису. Поставимо задачу шляхом переходу в новий простір на основі побудови відображення за рахунок застосування ортогонального розкладання по системі ФУ сформувавши більш компактний скорочений опис, що має кращі властивості для класифікації.

Розглянемо перетворення, пов'язане з розкладанням по сімейству прямокутних базисних функцій, в якості яких можна застосувати дискретні ФУ. Тепер перетворимо опис шляхом множення матриць:

$$U = Z * A, \quad (2.2)$$

де  $U$  – матриця Адамара, що включає вектори ФУ  $w_1, \dots, w_n$ .

Вираз (2.2) можна трактувати як одномірне перетворення Уолша, виконане для кожного з векторів опису.

Метод синтезу ефективного інформативного опису через відображення  $\Theta: Z \rightarrow U$  представимо у вигляді послідовності кроків:

1. Застосовуючи до вихідного опису перетворення (2.2), отримаємо матрицю  $U = \{u_{ij}\}$  розміром  $s \times n$ . Обчислимо квадрати елементів і перейдемо до аналізу матриці енергій  $U^2 = \{u_{ij}^2\}$ . Значення  $u_{ij}^2$  відповідає енергії ХО з номером з опису  $Z$ , що припадає на базисний вектор  $w_j$  (з номером  $j$ ).

Як відомо, при поданні до ортонормального базису повна енергія сигналу дорівнює сумі компонентів енергії, які відповідають усім базисним векторам.

2. З метою побудови інформативного кортежу для отриманого спектра здійснимо аналіз за допомогою дисперсійного критерію, обчислюючи дисперсії елементів в кожному

зі стовпців матриці  $U^2$ :  $\sigma_i^2 = \frac{1}{s-1} \sum_{k=1}^s (u_{ik}^2 - \bar{u}_i^2)^2$ , де  $\bar{u}_i^2 = \frac{1}{s} \sum_{k=1}^s u_{ik}^2$  – середнє,  $i = \overline{1, n}$ .

Значення дисперсії характеризують розкид відгуків ХО для окремих ФУ спектра [29].

Далі відсортуємо отриманий масив дисперсій  $\sigma_1^2, \dots, \sigma_n^2$  зі збереженням номерів ФУ, для яких вони обчислені.

3. Виберемо фіксоване число  $m \ll n$  ФУ, яким відповідають найбільші дисперсії. Вибрані ФУ  $w_1, \dots, w_m$  з їх вихідними номерами  $j_1, \dots, j_m$  будемо використовувати як найбільш інформативну частину базису для отримання ефективного опису. Тут

застосований підхід, де з метою виділення та фільтрації протяжних об'єктів на зображенні використовується обмежений набір ортогональних функцій з найбільш відповідними властивостями.

4. Здійснимо класифікацію векторів опису на класів відповідно до обраного інформативним кортежем ФУ. Для цього проаналізуємо обрані стовпців матриці  $U^2$  і для кожного з  $s$  векторів вихідного опису реалізуємо оптимальне правило  $R_Z$  віднесення вектора  $z_q \in Z$  до одного з класів  $k$  у діапазоні  $1, \dots, m$ :

$$R_Z : z_q \in k \mid \arg \max_{h=1, \dots, m} u_{qh}^2 = k, k=1, \overline{m}, q=1, \overline{s}. \quad (2.3)$$

Таким чином, кожен елемент вихідного опису буде віднесений до класу, для якого квадрат відгуку з відповідною функцією Уолша буде максимальним. Правило  $R_Z$  реалізує класифікацію на основі оптимальної узгодженої фільтрації елементів ОО в ортогональному базисі кортежу ФУ. Тут можливі інші варіанти обробки з обмеженнями і відсіювання, коли вектор ХО не класифікують ні до одного з класів в силу незначущості його відгуку [18].

5. В результаті здійснення класифікації (2.3) отримаємо розподіл  $s$  елементів опису у вигляді МХО за класами. Сформуємо стислий опис в агрегованому вигляді як вектор цілих чисел

$$u^* = (u_1, \dots, u_m), \sum_{j=1}^m u_j = s, \quad (2.4)$$

елементи  $u_j$  якого – кількості векторів ХО, віднесені до класу  $j$ .

З огляду на те, що еталони класів можуть містити в своєму складі різну кількість ХО, на практиці часто необхідно здійснити нормування опису (2.4):  $u_i = u_i / s$ , що гарантує виконання співвідношення  $\sum_{j=1}^m u_j = 1$ .

В результаті послідовність перетворення вихідного опису в процесі стиснення можна формалізувати як:

$$Z \Rightarrow U \Rightarrow u^*, \quad (2.5)$$

або у функціональному вигляді:

$$u^* = F_2[F_1[Z]], \quad (2.6)$$

де  $F_1$  – перетворення до простору ФУ;

$F_2$  – класифікація векторів в цьому просторі.

Запропонований метод перетворює вихідний опис у вигляді матриці розміру  $S \times N$  до вектору з числових значень, причому можна вибрати  $m \ll n$ , що істотно знижує обчислювальні витрати особливо при великих значеннях числа векторів  $S$ , яке в практичних завданнях іноді досягає 200-1000. Важливою перевагою при цьому є те, що перетворення еталонних ОО до стиснення виконується один раз на попередньому етапі і в цілому не впливає на загальний час класифікації. Відкидання частини коефіцієнтів розкладання, як показали дослідження, не викликає істотних помилок або втрати потрібних властивостей. Розміром обсягу опису (параметром) можна управляти, підвищуючи характеристику достовірності [15].

Вибір ФУ як ортогонального базису, крім простої апаратної реалізації, додатково можна обґрунтувати їх високими фільтраційними властивостями щодо дії адитивних флуктуаційних перешкод, які характерні для прикладних задач обробки зображень.

Важливим окремим випадком розглянутого методу є варіант  $m = n$ , тобто ситуація, якщо всі наявні ФУ беруть участь в класифікації. Для цього варіанту стиснення опису можливо на основі аналізу значимості елементів уявлення (2.6) шляхом відкидання малих або виділення найбільших за величиною значень серед  $u_j$ .

Дослідження показують, що енергія спектра при поданні в базисі ФУ для векторів ХО на практиці розподіляється дуже нерівномірно. Три ФУ в значній мірі домінують над іншими, що на практиці призводить до класифікації переважної більшості ХО до цих трьох класів. За рахунок цього фактора відбувається надмірне стиснення інформації, і в результаті в цілому погіршуються показники якості класифікації, особливо при дії перешкод. Така картина спостерігається для декількох досліджуваних баз зображень, що говорить про об'єктивність зазначених спектральних властивостей і отриманих висновків [40].

Використаємо стислий опис виду (2.6) для класифікації шляхом пошуку еталона з найбільшим значенням міри подібності за правилом:

$$c(u) = \arg \operatorname{opt}_{q \in \{1, \dots, KQ\}} \mathcal{Q}[u, u(q)], \quad (2.7)$$

де  $c(u)$  – функція класифікації опису  $u$ ,

$\mathcal{D}[u, u(q)]$  – міра подібності між векторами в просторі  $\mathbb{R}^m$ ,

$u(q)$  – вектор опису зразка з номером  $q$ ,

$K_Q$  – множина класів.

У результаті подання з урахуванням стиснення формується множина  $K_m$  класів ознак, яке включає  $m$  класів і засноване на виділеному кортежі ФУ. При цьому в процесі класифікації опис об'єкта будується з урахуванням ФУ, що в результаті забезпечує додаткові узгодження і перешкодозахищеність.

За рахунок подання координатного векторного простору ХО в ортогональному базисі ФУ отримуємо новий векторний простір, переходячи від уявлення об'єкта у вигляді множини векторів до одного вектору, що відбиває спектральне подання для найбільш значимих (інформативних) елементів базису. Тепер класифікація зводиться до подання об'єкта у вигляді вектора  $u$  з заданим для бази множиною класів.

Міра подібності  $\mathcal{D}$  в (2.7) може бути обчислена або шляхом голосування, або шляхом векторного представлення об'єкту і зіставлення описів як векторів. Зрозуміло, що опис може при цьому містити елементи перешкод і помилкових фонових об'єктів, що призводить до викривлення вектора (3.3). Зокрема, можна використовувати метрику в просторі мультимножин:

$$\rho(a, b) = \frac{\sum_i w_i |u_a(i) - u_b(i)|}{\sum_i w_i \max[u_a(i), u_b(i)]}, \quad (2.8)$$

де  $a, b$  – зіставляються вектори;

$u_a(i), u_b(i)$  – значення  $i$ -х компонентів векторів;

$w_i$  – вагові коефіцієнти.

Крім властивостей розрізнення і перешкодозахищеності, важливим моментом залишається оцінка виграшу у швидкодії запропонованого методу в порівнянні з методом традиційного голосування на МХО. Проведемо її шляхом підрахунку числа обчислювальних операцій.

Позначимо  $t_E$  – час обчислення квадрата евклідової відстані або скалярного множення між векторами ХО розміром  $n$ ,  $t_c$  – час порівняння і підсумовування двох чисел при пошуку мінімуму відстані і інкрементація числа голосів,  $s, s_1$  – потужності



описів еталона і об'єкта. Тоді оцінка часу при обчисленні подібності у вигляді числа голосів виглядає як  $T_v = ss_1(t_E + t_c)$ . Час для обчислення подібності відповідно до обробки (2.2) – (2.4) можна оцінити як  $T_m = s_1[m(t_E + t_c) + t_c] + (t_E m) / n$ , де вираз  $(t_E m) / n$  відповідає обчисленню подібності між векторами розміром  $m$ . Ставлення оцінок має вигляд:

$$\alpha = \frac{T_v}{T_m} = \frac{ss_1(t_E + t_c)}{s_1[m(t_E + t_c) + t_c] + (t_E m) / n}. \quad (2.9)$$

Детальний аналіз виразу (2.9) показує, що найбільшою мірою величина  $\alpha$  залежить від співвідношення  $s / m$ , тобто паралельно збільшується зі збільшенням  $s$  і зменшується зі збільшенням  $m$ . Відсутність залежності від  $n$  можна пояснити тим, що обидва варіанти включають аналіз всього опису об'єкта, що включає  $s_1$  векторів однакової розмірності  $n$  [24].

Розрахунки показують, що конкретно при  $m = 8$ ,  $s = 100$  значення  $\alpha \approx 12,5$ , а при  $s = 200$  –  $\alpha \approx 25$ . Як бачимо, переваги запропонованого методу в швидкодії для практичних значень параметрів можуть досягати 10-20 разів.

Проведені експерименти з використанням програмних моделей показали більш високу ефективність при класифікації із застосуванням запропонованих перетворень не тільки в плані швидкодії, але і за рівнем перешкодозахищеності. До рівня шуму  $\sigma^2 = 0,02$  ймовірність правильного розпізнавання дорівнює 0,94, далі, зі збільшенням дисперсії шуму  $\sigma^2$  крива ймовірності для розробленого методу вище, і при  $\sigma^2 = 0,03$  ймовірність становить величину 0,8, в той час як традиційний метод голосування забезпечує ймовірність правильної класифікації лише на рівні 0,6.

## 2.2 Дослідження властивостей трансформованих описів

Метод SURF формує опис  $Z \subset \mathbb{R}_1^n$  з множини  $\mathbb{R}_1^n$  векторів  $z = (z_1, \dots, z_n)$ ,  $z \in Z$  векторного простору  $\mathbb{R}^n$ ,  $\mathbb{R}_1^n \subset \mathbb{R}^n$ .

Множина  $\mathbb{R}_1^n$  визначено у вигляді множини всіх  $n$ -мірних векторів з речовими компонентами, а евклидова норма векторів з  $\mathbb{R}_1^n$  дорівнює 1:

$$\|z\| = \sqrt{\sum_{i=1}^n z_i^2} = 1. \quad (2.10)$$

Лінійним перетворенням  $L$  вектора  $z \in \mathbb{R}_1^n$  до вектору  $u = (u_1, \dots, u_n)$  з нової множини  $U_1^n$  векторів  $U_1^n \subset \mathbb{R}^n$ , тобто  $L: \mathbb{R}_1^n \rightarrow U_1^n$ .

Запишемо це перетворення в розгорнутому вигляді:

$$U = \begin{bmatrix} u_1^1 \dots u_n^1 \\ u_1^2 \dots u_n^2 \\ \dots \\ u_1^s \dots u_n^s \end{bmatrix} = ZA = \begin{bmatrix} z_1^1 \dots z_n^1 \\ z_1^2 \dots z_n^2 \\ \dots \\ z_1^s \dots z_n^s \end{bmatrix} \times \frac{1}{n} \begin{bmatrix} a_{11} \dots a_{1n} \\ a_{21} \dots a_{2n} \\ \dots \\ a_{n1} \dots a_{nn} \end{bmatrix}. \quad (2.11)$$

Вираз (2.11) реалізується шляхом множення (рядок на стовпець) матриці вихідного опису  $Z$ , що включає  $s$  векторів  $XO$  у вигляді рядків, на ортогональну матрицю Адамара  $A$  розміром  $n \times n$  з нормувальним коефіцієнтом  $1/n$ .

Вираз  $U = ZA$  через симетричність можна вважати узагальненим записом традиційного для алгебри лінійного перетворення виду  $u = Az$ , де  $u \in \mathbb{R}^n$ ,  $z \in \mathbb{R}^n$  –  $n$ -мірні вектора-стовпці, тому що за будь-яку зміну базису, пов'язану з переходом до нового базису, відповідає невироджене лінійне перетворення координат кожного вектора. На рисунку 2.1 наведено приклад дескриптора, який демонструє його перетворення шляхом множення на матрицю Адамара.

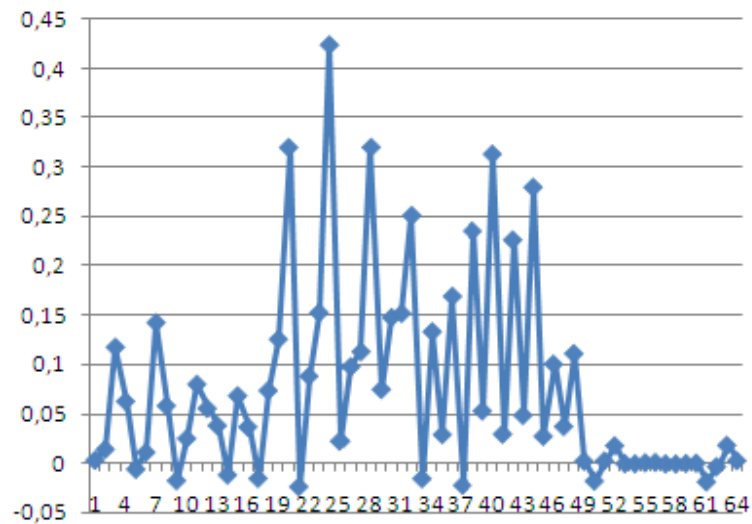


Рисунок 2.1 – Дескриптор ХО, сформований методом SURF

Матриця  $A$  є матрицею лінійного перетворення в координатному базисі векторів з  $\mathbb{R}_1^n$ . Кожне невироджене лінійне перетворення, як відомо, має єдине зворотне лінійне перетворення, матриця якого є зворотною до матриці  $A$ . Для перетворення (2.10) вихідне опис може бути отримано з  $U$  відповідно до вираження  $Z = UA^{-1}$  шляхом множення на зворотну матрицю. З огляду на, що для ПУА зворотня матриця через властивості ортогональності збігається з  $A$  за винятком коефіцієнта  $\frac{1}{n}$ , то опис  $Z$  може бути обчислено за формулою зворотного перетворення  $Z = UA$ .

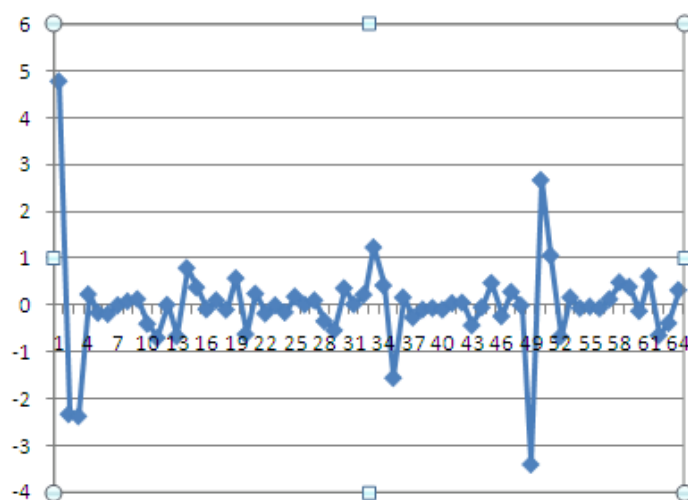


Рисунок 2.2 – Дескриптор ХО після перетворення Уолша

Для конкретного вектора  $z \in \mathbb{R}_1^n$  взаємозв'язок з його спектром  $u \in U_1^n$  відповідно до (2.10) визначається парою дискретних ПУА [11]:

$$\begin{cases} u_i = \frac{1}{n} \sum_{k=1}^n z_k a_{ki}, \\ z_i = \sum_{k=1}^n u_k a_{ki} \end{cases}, \quad (2.12)$$

де  $a_{ki}$  – дискретна ФУ (стовпець матриці  $A$ ) з порядковим номером  $i$ .

Традиційно в теорії обробки сигналів відліки векторів змінюються в межах від 0 до  $n-1$ .

Ортогональна матриця – це квадратна матриця з елементами, для якої її зворотна матриця дорівнює транспонованій, тобто  $A^{-1} = A^T$ .

Для матриці Адамара виконується співвідношення:

$$AA^T = nE,$$

де  $E$  – одинична матриця розміру  $n$ . Стовпці і рядки ортогональної матриці утворюють системи ортонормованих векторів, для яких скалярний добуток рядка (стовпчика) на себе дорівнює 1, а на будь-яку іншу – 0. Ортогональні матриці відповідають лінійним операторам, що переводять ортонормований базис лінійного простору в ортонормований.

Строго в математичному плані визначення ортонормированного ПУА записується у вигляді наступного співвідношення [20]:

$$\begin{cases} u_i = \frac{1}{\sqrt{2^r}} \sum_{k=1}^n z_k a_{ki} \\ z_i = \frac{1}{\sqrt{2^r}} \sum_{k=1}^n u_k a_{ki} \end{cases}, \quad (2.13)$$

яке відрізняється від (2.12) нормувальними коефіцієнтами.

Значення  $r = \log_2 n$  найбільш просто визначається в разі, якщо  $n$  є ступінь двійки.

Для цього випадку матриця Адамара може бути побудована за загальною формулою:

$$A_{2n} = \frac{1}{\sqrt{2}} \begin{bmatrix} A_n & A_n \\ A_n & -A_n \end{bmatrix}, \text{ починаючи з } A_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}. \quad (2.14)$$

Конкретно, наприклад, при  $n=64$  значення  $\sqrt{2^r}$  дорівнює 8, а при  $n=128$  величина  $\sqrt{2^r} = 8\sqrt{2}$ . ПУА у вигляді (2.14) зберігає скалярний добуток між векторами, тобто виконана умова  $(Ax, Ay) = (x, y)$ , де  $(x, y)$  – скалярний добуток,  $x \in \mathbb{R}_1^n, y \in \mathbb{R}_1^n, Ax \in U_1^n, Ay \in U_1^n$ .

Іншими словами, норма і кут між векторами зберігаються, тобто  $\|Ax\|^2 = \|x\|^2$  та  $\cos(Ax \square Ay) = \cos(x \square y)$ .

Це дозволяє зробити висновок, що відображення  $L: \mathbb{R}_1^n \rightarrow U_1^n$  в ситуації (2.14) переводить простір в себе, тобто  $U_1^n = \mathbb{R}_1^n$ . При цьому  $U_1^n$  так само, як і, є множина векторів на багатовимірній одиничній кулі, тобто з нормою 1.

У практичному плані це дає можливість використовувати одні і ті ж значення параметра порогу при встановленні еквівалентності в просторах вихідних і трансформованих векторів. У той же час недоліком (2.13) в порівнянні з (2.14) можна вважати необхідність виконання додаткової операції ділення для кожного з елементів.

В інтерпретації теорії фільтрації сигналів подання до базису Уолша можна розглядати як застосування до вихідного вектору ХО кінцевого банку фільтрів, що відображає спектральні властивості вектора опису. Побудова скороченого базису на основі використання неповного кортежу ФУ, відібраного за деяким критерієм, дозволяє отримати компресійне (компактне) уявлення, що володіє високими показниками швидкодії при розпізнаванні [26].

Важливим моментом є дослідження і підтвердження інваріантних властивостей матриці  $U$ , тому опис  $Z$ , сформований детектором SURF для об'єкта, є множина векторів, що володіють властивістю інваріантності до геометричних перетворень аналізованого об'єкта з певної групи перетворень.

З огляду на вже проаналізовані особливості ПУА, в цілях успішного використання при розпізнаванні необхідно визначити дотримання властивості інваріантності для перетвореного опису  $U$ , тобто перевірити виконання умови  $\forall g \in G: U(B) = U(gB)$ . Проведені експериментальні дослідження підтвердили цей факт.

## 2.3 Компресування опису візуальних об'єктів шляхом фільтрації його компонент за критерієм стабільності

Механізм людського зору в порівнянні із зоровими системами штучного інтелекту у вищій мірі володіє можливістю виділяти найбільш значущу для розпізнавання відеоінформацію. Це можуть бути як локальні характеристики (ключові точки), так і глобальні дані (взаємне розміщення частин об'єкта, контраст, відносні розміри і ін.).

Наприклад, в процесі спостереження за об'єктом, що рухається людське око спирається, як правило, на невелике число стійких до змін характеристик об'єкта.

Природно висунути біонічну гіпотезу про те, що застосування розроблених штучних інтелектуальних методів має бути пов'язане з формуванням незначного по потужності, але в той же час інформативного в плані достатності для забезпечення процесу розпізнавання підмножини стійких за деяким критерієм ознак.

Формально зниження розмірності простору ознак в загальному випадку зводиться до побудови відображення:

$$\Xi: X_1 \times X_2 \times \dots \times X_p \rightarrow Y_1 \times Y_2 \times \dots \times Y_q, \quad q < p, \quad (2.15)$$

де  $X_1, \dots, X_p$  – початковий набір ознак,

$Y_1, \dots, Y_q$  – новий набір ознак,

$\times$  – символ декартова множення,

$p, q$  – розмірності вихідного і стисненого просторів.

При цьому вважається, що кожна з ознак має свою область значень. Стосовно до описів  $X_i \in \mathbb{R}^n, Y_j \in \mathbb{R}^m$  можна розглянути два основних способи побудови відображення стиснення: шляхом скорочення числа векторів в множини  $Z$  і / або зменшення розмірності окремого вектора, тобто через перехід до простору  $\mathbb{R}^v, v < n$ . Зокрема, побудова  $\mathbb{R}^v$  можна реалізувати шляхом переквантування або передискретизації дескриптора ознак за значенням окремої компоненти або за загальною кількістю компонент [33].

Скорочення числа векторів в описі  $Z$  передбачає побудову стисненого опису  $Z^*$  на основі відображення вигляду  $\Xi: Z \rightarrow Z^*$ , варіантом якого є формування підмножини  $Z^* \subset Z$  значно меншої потужності шляхом застосування деякої процедури фільтрації. У

теорії розпізнавання цей спосіб асоціюється з формуванням множини найбільш значущих ознак. Інший варіант побудови полягає в смислового агрегування елементів  $Z$  шляхом об'єднання їх властивостей в більш глобальні структури даних, що формують  $Z^*$ .

В цілому застосування стиснення опису направлено на зменшення обсягу обчислень і зниження вимог до застосовуваних методів розпізнавання з точки зору швидкодії. Ідея пропонує методів полягає в побудові ефективної процедури для пошуку незначної кількості «хороших» особливостей зображення.

Виграш у часі розпізнавання пропорційний відношенню потужностей старої і нової (стислої) систем ознак. У той же час треба мати на увазі, що зменшення обсягу опису шляхом його стиснення може привести до зниження ймовірності розпізнавання в умовах перешкод.

Візьмемо за основу зниження розмірності простору ознак одне з найважливіших властивостей – стійкість значень інваріантних ознак МХО при дії геометричних перетворень. Зрозуміло, що стабільність в повній мірі визначається використовуваним детектором ознак. Інші принципи фільтрації даних можуть застосовувати, наприклад, відбір найбільш інформативних елементів опису. Зауважимо при цьому, що традиційні детектори характерних ознак SIFT, SURF частково вже реалізують функцію відбору найбільш значущих для розпізнавання компонентів опису [40].

Побудуємо спочатку множину зображень, яка буде основою встановлення та формування стійкої скороченої підмножини ознак. Для цього виберемо довільний еталон з відео-бази і піддамо його фіксованому числу  $N$  різних значень геометричних перетворень з групи  $G = \{g_i\}_{i=1}^N$ . Одним з варіантів обробки є дискретизація в практичних додатках простору перетворень.

Наприклад, для перетворення повороту послідовно обертаємо еталон на дискретні кути  $g_1 = 20^\circ$ ,  $g_2 = 40^\circ$ , до  $g_{18} = 360^\circ$ , де  $N = 18$ . Ще одним прийнятним способом формування  $G$  може бути процедура випадкового вибору його  $N$  елементів по рівномірному закону з деякої кінцевої множини.

Для кожного з перетворених зображень шляхом застосування детектора формуємо опис  $Z_1$  у вигляді МХО.

Позначимо  $s$  – розмір вихідного еталонного опису  $Z$  (при куті повороту  $g_0 = 0^\circ$ ),  $s_1$  – розмір опису  $Z_1$  перетвореного еталона.

Застосуємо тепер відображення  $\Theta$  встановлення відповідностей  $\Theta: Z \rightarrow Z_1$  елементів двох описів, в процесі реалізації  $\Theta$  для кожного  $z \in Z$  спробуємо встановити елемент  $z_1 \in Z_1$  через побудову відповідності  $\beta(z, z_1) \in \{0, 1\}$ , яке дорівнює 1, якщо  $z_1$

сформован, та 0 – в іншому випадку. Формально відповідність  $\Theta$  між множинами  $Z, Z_1$  з елементами  $z \in Z, z_1 \in Z_1$  визначено як підмножина декартової множини  $\Theta \subseteq Z \times Z_1$  і може бути задано таблицею.

З кожним з сформованих відповідностей  $\beta(z, z_1)$  зв'яжемо значення  $\rho(z, z_1)$  метрики, шляхом мінімізації значень якої воно синтезується. В якості  $\rho(\cdot)$  може бути використана довільна метрика в  $\mathbb{R}^n$ . Виконаємо відображення  $\Theta$ , в результаті для кожного  $z \in Z$  отримаємо значення  $\beta$ . Сформуємо бінарний масив  $\beta^g = \{\beta_1^g, \beta_2^g, \dots, \beta_s^g\}$ , який зв'яжемо зі значенням перетворення  $g$ .

Проведемо випробування для всіх зображень, отриманих шляхом застосування до вибраного еталону геометричних перетворень з  $G = \{g_i\}_{i=1}^N$ . Для кожного номера  $q$  елемента у векторі підрахуємо число сформованих відповідностей

$$k_q = \sum_{i=1}^N \beta_q^i, \quad q \in [1, s], \quad k_q \in [0, N], \quad (2.16)$$

в серії з  $N$  випробувань. Будемо використовувати критерій (2.16) як характеристику стійкості  $q$ -го ознаки еталона. За результатами обробки на множини  $G$  сформуємо таблицю значеню критерію  $k_q$  для всіх  $s$  ознак.

Для однозначних відповідностей виконується нерівність  $k_q \leq N$ . При цьому факт можливого формування відповідності з «чужою» ознакою не контролюється. Правильність встановлення відповідності повністю визначається відображенням  $\Theta$ . Значення критерію  $k_q$  відображає рейтинг стійкості для  $q$ -ої ознаки на навчальній множині.

Далі опрацюємо елементи множини  $\{k_q\}_{q=1}^s$  шляхом ранжування і виділимо в ньому підмножину  $Q^m$  з  $m$  номерів елементів з найбільшим значенням рейтингу  $k_q$ . Ще одним варіантом обробки може бути формування  $Q^m$  з номерів елементів, що мають максимальне значення  $k_q$ .

У цьому випадку величина виходить як результат аналізу. Зрозуміло, що вибір  $m$  для успішного вирішення задачі розпізнавання залежить від еталонів. Описаний підхід до формування стисненого опису (схема на рисунку 2.3) можна характеризувати як синтез підмножини найбільш стійких компонентів опису [32].



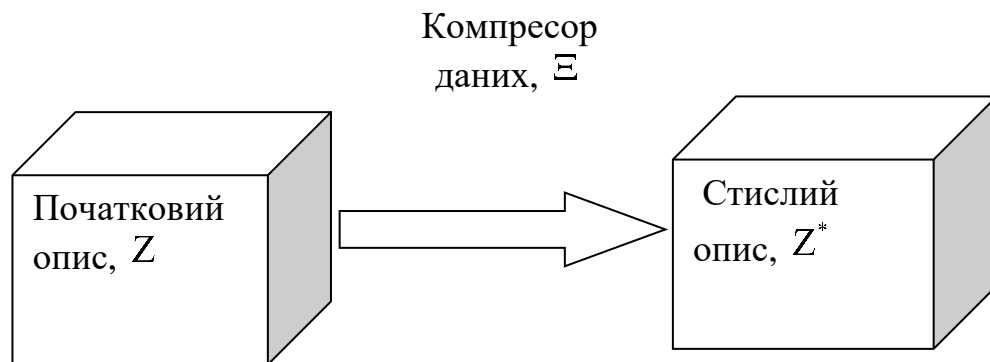


Рисунок 2.3 – Схема компресування

Іншим варіантом способу компресування даних може бути використання для рейтингування компонентів критерію у вигляді суми значень мінімумів відстаней, при яких сформовано кожна відповідність з  $Q^m$ . Значення  $r_q$  цього критерію опишемо як суму:

$$r_q = \sum_{i=1}^{k_q} \rho(z_q, z_{li}), \quad (2.17)$$

де  $\rho(z_q, z_{li})$  – відстань для сформованої відповідності  $\beta_i^q = 1$ .

Часто в якості  $\rho$  використовують евклідову метрику. Чим менше значення  $r_q$ , тим кращим чином сформовано стисла множина  $Q^m$ . В ідеальному випадку  $r_q = 0$ . Можливо також спільне використання критеріїв (2.16), (2.17) в одній процедурі. Значення критеріїв представимо у вигляді таблиці 3.1.

Обробка значень таблиці 2.1 з метою формування  $Q^m$  може бути виконана кількома способами, такими як:

а) порогова обробка, при якій до складу  $Q^m$  відбираються ті номери ознак, для яких  $k_q \geq \delta_k$ , де  $\delta_k$  – поріг для числа;

б) ранжування  $\{k_q\}$  по спадаючій і виділення з отриманого ряду  $m$  перших ознак.

Таблиця 2.1 – Значення критеріїв стійкості ознак

Номер ознаки еталона	1	2	...		s-1	s
Число відповідностей	$k_1$	$k_2$	...		$k_{s-1}$	$k_s$
Сума відстаней	$r_1$	$r_2$	...		$r_{s-1}$	$r_s$

У будь-якому випадку для еталона має бути сформовано не менше ніж  $m$  ознак, з використанням яких буде здійснюватися розпізнавання або класифікація.

Результат формування стисненого опису  $Q^m$  на основі зіставлення залежить від ряду факторів, зокрема, від способу отримання ознак (більшою мірою) і від правила, що встановлює еквівалентність компонентів опису. Властивості детектора і правило визначаються рядом параметрів. До них відносяться: поріг  $\delta_\rho$  на величину мінімуму відстані  $\rho(Z_q, Z_{i1})$  між векторами (в традиційному варіанті прийнято  $\delta_\rho=0,24$ , що відповідає 1,5% від максимально можливого відхилення [53]), значення відносини мінімуму до найближчого мінімуму (приймається рівним 0,7). Крім того, на результат в значній мірі впливає також спосіб моделювання геометричних перетворень. Так, при застосуванні методів інтерполяції підвищується число досягнутих відповідностей за рахунок зростання числа точок [31].

Розглянуту обробку можна застосувати і для інших перетворень, наприклад, масштабування, а також для перетворень комплексованих типів. У той же час для найбільш простих перетворень зсуву на цілі величини, де немає спотворень, пов'язаних з дискретизацією растра, обговорюваний підхід прямо не застосуємо, тому що внаслідок відсутності втрат точок опису значення  $K_q$  для всіх ознак дорівнюватиме  $N$ , і стабільні точки можна вибирати довільним чином. Для змішаних перетворень, що застосовуються на практиці (зміщення, поворот, масштаб) практичне застосування підходу не викликає ускладнень.

За результатами узагальненого аналізу для всіх еталонів відео-бази сформуємо підмножини  $Q_i^m$  для  $i$ -го еталона.

В результаті досягаємо скорочення обсягу інформації в базі даних, представляючи кожен з еталонів описом з  $m$  елементів.

Ця процедура допускає узагальнення. Наприклад, можна встановити різну кількість найбільш стійких характерних ознак для різних еталонів порогу для значення метрики. Інший варіант побудови стислій бази може бути пов'язаний з послідовною процедури відбору ознак, які не містяться в уже синтезованих описах інших еталонів.

Відображення (2.17) було реалізовано програмно шляхом формування для еталонного опису масиву «стабільних» точок при здійсненні кінцевого числа  $N=18$  поворотів. Множина стабільних точок формувалося як набір з найбільшим числом повторень. На рисунок 2.4 наведені приклади еталонного зображення шахової фігури коня зі сформованими характерними ознаками, а також перетворене зображення з різним

числом стабільних точок при значеннях параметра  $\delta_\rho=0,24$  (64 точки),  $\delta_\rho=0,16$  (31 точка) [14].

При  $\delta_\rho=0,08$  число стабільних точок дорівнювало 6. Такі значення порогу відображають точність зіставлення у вигляді відсотка від максимального значення відхилення між векторами SURF в евклідовій метриці: 0,08 – 0,5%, 0,16 – 1%, 0,24 – 1,5%.

З рисунок 2.4 видно, що шляхом зміни параметра  $\delta_\rho$ , фактично відбиваючи розмір допустимої околиці для еквівалентних векторів в  $R^n$ , можна управляти числом стабільних точок.

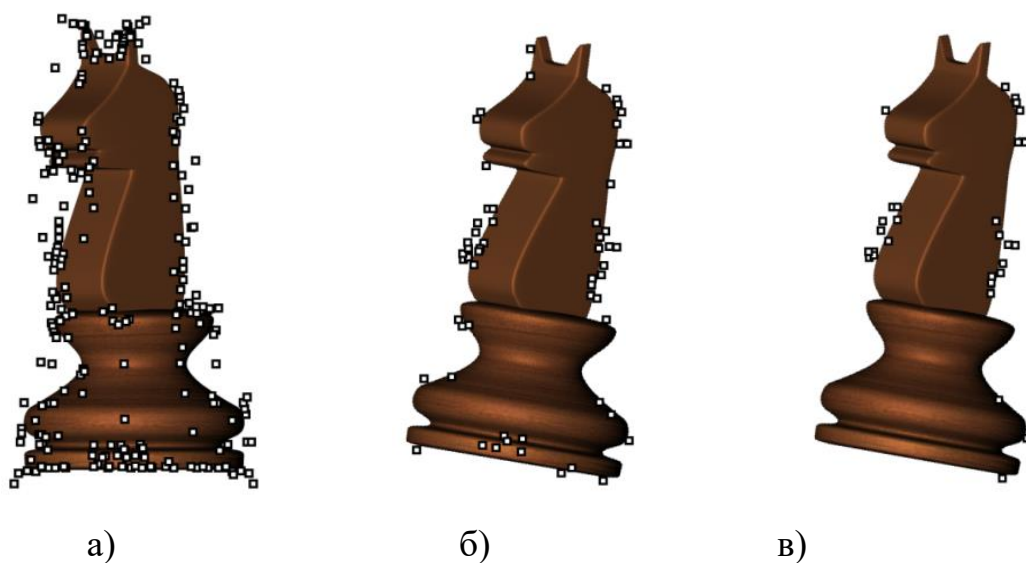


Рисунок 2.4 – Зображення з ключовими ознаками: а - еталонне (число ознак 239), б – при  $\delta_\rho=0,24$  (64), в –  $\delta_\rho=0,16$  (31)

На рисунку 2.4 наведено зображення з виділеними найбільш стійкими точками при повороті і масштабі.

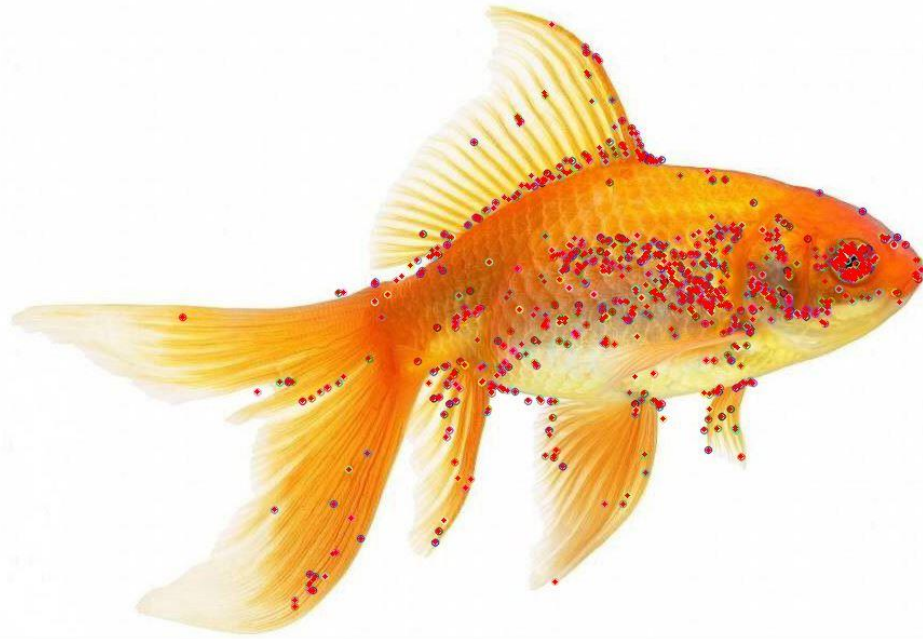


Рисунок 2.5 – Зображення об'єкта з зазначеними найбільш значущими точками при повороті і масштабі

Таким чином, необхідно відзначити, що компресування – це потужний засіб підвищення ефективності розпізнавання. В першу чергу слід відзначити, що безпомилкова класифікація зберігається на рівні до 40% від усього числа точок опису, що відповідає допустимому рівню компресування в 2,5 рази.

## 2.4 Оптимізація параметра порога при зіставленні описів об'єктів

Задачі багатокритеріальної оптимізації знаходять широке застосування в практиці штучного інтелекту, зокрема, при вирішенні задач багатофакторного оцінювання при управлінні інформаційно-технологічними системами. Наприклад, з використанням теорії багатокритеріальної оптимізації з'являється можливість оптимального вибору значень порогів, використовуваних при зіставленні МХО.

В даному випадку визначення оптимального значення параметра полягає в пошуку екстремуму функції двох змінних.

Особливістю даного завдання, як і всіх багатокритеріальних задач прийняття рішень, є те, що використовувані критерії, як правило, є суперечливими і різномірними. Хоча всі застосовувані критерії мають цілком ясне об'єктивний зміст, а оцінки за критеріями – кількісне вираження, часто важко отримати єдину кількісну модель, що описує проблему в цілому.

В даному випадку необхідно вибрати ту з заданих альтернатив (значень порогу), де досягається найкращий компроміс між критеріями.

В процесі детального дослідження властивостей детектора SURF з'ясувалося, що якість класифікації істотно залежить від порогового значення  $\delta_z$ , визначального еквівалентність двох елементів описів об'єктів, представлених у вигляді множини векторів. Якщо відстань між елементами  $z_1, z_2$  менше порогу, тобто  $\rho(z_1, z_2) \leq \delta_z$ , то компоненти  $z_1, z_2$  вважаються еквівалентними між собою, що безпосередньо впливає на результат класифікації.

На основі значення  $\delta_z$  визначається величина подібності між описами як число голосів, відданих елементами опису еталона за розпізнавання об'єкту. Для конкретної бази риб допустимий діапазон порогових значень склав відрізок  $[0,08; 0,8]$ . при порозі  $\delta_z = 0,8$  в процесі порівняння описів визначаються всі відповідності точок перетвореного зображення ідеалу, що характеризує виконання властивості інваріантності в повному обсязі, тобто для всіх точок опису. При значенні порога  $\delta_z < 0,8$  лише для частини точок опису дотримується ця властивість. При цьому для крайньої ситуації  $\delta_z = 0,08$  число інваріантних точок не перевищує 5% від загального обсягу опису зразка, що може викликати проблеми розрізнення об'єктів між собою. Поставимо задачу вибору оптимального значення порогу  $\delta_z \in [0,08; 0,8]$ , що забезпечує достовірну оцінку подібності описів при збереженні розрізнення між об'єктами різних класів [25].

Основою для оцінювання подібності описів, зокрема, при формуванні множини стійких точок, є матриця міжкласових відстаней  $H = \{h_{ij}\}$ , в якій елементи головної діагоналі  $h_{ii}$  – це відстані між еталоном і перетвореним еталоном, а елементи  $h_{ij}$  – відстані між перетвореним еталоном і іншими еталонами бази. Результати та якість класифікації повністю визначаються значеннями  $H$ .

Ідеальним випадком є ситуація, коли діагональні елементи містять максимально можливу кількість голосів, яке збігається з кількістю інваріантних точок в описі зразка. При цьому інші елементи матриці повинні бути близькі до 0. Це показує, що всі зображення максимально помітні між собою. Однак при зміні значень порога  $\delta_z$  елементи  $h_{ij}$  збільшуються, що показує зростання подібності між різними стандартами. Таким чином, ситуація породжує необхідність оптимізації значень двох часткових критеріїв: необхідність збереження якомога більших значень діагональних елементів  $h_{ii}$  при одночасному зниженні значень інших елементів  $h_{ij}$ .

Формалізуємо це у вигляді критеріїв  $A_1(\delta_z)$ , який потрібно максимізувати, і  $A_2(\delta_z)$ , який необхідно мінімізувати:

$$\max A_1(\delta_z) = \frac{1}{c} \sum_{i=1}^c h_{ii} / \max(h_i^*) \quad (2.18)$$

$$\min A_2(\delta_z) = \frac{1}{c} \sum_{i=1}^c 1(\max_{j=1, \dots, c-1} h_{ij} \geq 0,5h_{ii}) \quad (2.19)$$

де  $A_1$  – середнє по числу еталонів ставлення діагональних елементів  $H$  до їх максимально можливих значення  $\max(h_i^*)$ ,

$A_2$  – частка числа рядків матриці  $H$ , в яких найближчий до діагонального значенням максимум більше половини діагонального елемента  $h_{ii}$ ,

$c$  – число еталонів бази,

а функція  $1(\dots)$  для рядка з  $H$  в (3.14) визначається як

$$1(\max_{j=1, \dots, c-1} h_{ij} \geq 0,5h_{ii}) = \begin{cases} 1, & \max_{j=1, \dots, c-1} h_{ij} \geq 0,5h_{ii}, \\ 0, & \max_{j=1, \dots, c-1} h_{ij} < 0,5h_{ii} \end{cases} \quad (2.20)$$

Критерій  $A_1$  відображає необхідність збереження максимально великого числа голосів, відданих елементами опису еталона за свій клас. Критерій  $A_2$  при цьому забезпечує мінімізацію подібності між різними еталонами бази (міжкласовий розкид). Таким чином, достовірне оцінювання подібності описів представимо як завдання двухкритеріальної оптимізації з двома суперечливими частковими критеріями (2.19), (2.20).

Для вирішення задач багатокритеріальної оптимізації існує багато підходів. У нашому випадку критерії різноспрямовані, але рівноправні, тому виключається використання рангових методів та прийняття умовно-оптимального рішення.

Для знаходження оптимального рішення в даній ситуації раціональним є застосування функції корисності, тобто використання згортки критеріїв, що забезпечує оптимізацію функції корисності. Зведемо двухкритеріальне завдання до нового агрегованого критерію шляхом згортки критеріїв  $A_2, A_1$  в один комплексний у вигляді:  $A = \min(A_2 - A_1)$  враховуючи, що значення обох часткових критеріїв знаходяться в діапазоні. В результаті діапазон значень критерію  $A$  становить відрізок  $[-1, 1]$ .

Для визначення оптимального значення порогу чисельно розрахована матриця міжкласових відстаней. В ідеальному випадку, без перешкод, кількість голосів, відданих еталоном при порівнянні з собою, має розкид всередині бази риб в межах від 100 до 200 точок. Це означає, що кількість точок, виділених детекторами ХО, нерівномірно, при цьому виявлена особливість бази – два еталона найбільш схожих між собою, навіть при відсутності перешкод між ними визначається близько 20 схожих точок [19].

Аналіз експериментальних даних показав, що з точки зору максимізації голосів оптимальним значенням порога є 0,8 (максимум критерію  $A_1$ ).

З точки зору мінімізації  $A_2$  найкращим значенням порога є 0,08. Оптимальне ж значення агрегованого критерію  $A = \min(A_2 - A_1)$  досягається при значенні порога 0,32 (воно дорівнює  $A=-0,52$ ). При цьому загальне скорочення числа голосів за еталони становить близько 25%, максимальне співвідношення між максимальним значенням голосів, відданих за один еталон (елементи  $h_{ij}$ ), і голосами, відданими за свій еталон ( $h_{ii}$ ), становить 0,3 в загальному випадку. На рисунку 2.5 представлена графічна інтерпретація значень часткових критеріїв  $A_2, A_1$  та агрегованого критерію  $A$  в залежності від величини  $\delta_z$ .

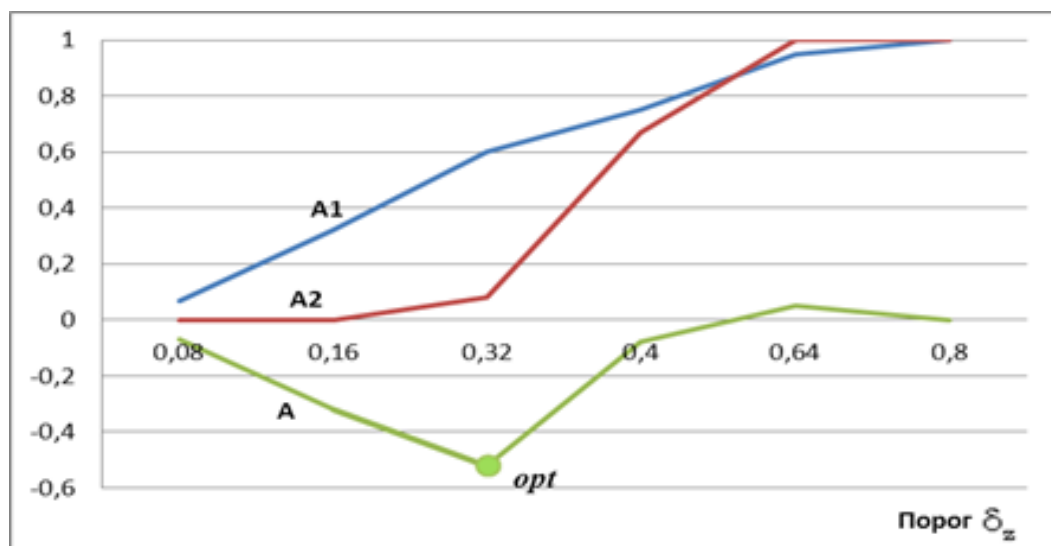


Рисунок 2.5 – Залежність значень критеріїв від значення порога  $\delta_z$

Запропонований метод побудови інформативних компактних описів на основі аналізу структурного уявлення в ортогональному базисі Уолша дозволяє здійснювати класифікацію об'єктів з високою швидкістю без зниження рівня перешкодозахищеності. Обчислювальні експерименти підтвердили припущення про можливість істотного

скорочення опису шляхом відкидання малих по величині спектральних коефіцієнтів перетворення.

При порівнянні рівня перешкодозахищеності методом голосування встановлено, що розроблений метод побудови компактних описів в просторі кортежу виділених ФУ не поступається по перешкодозахищеності методу голосування, що обумовлено інтегральними властивостями базису Уолша. Перевагою розробленого методу є значний вигаш у швидкодії: теоретично – в 10-20 разів, практично – в 30 разів у порівнянні з голосуванням.

Запропонована модель компресування інваріантних описів з використанням критерію стабільності їх значень на навчальній множині геометричних перетворень полягає в реалізації наступних кроків:

1) встановлення підмножини найбільш стабільних точок опису з точки зору збереження якості інваріантності до групи геометричних перетворень;

2) формування стисненого опису зі складу елементів підмножини стабільних компонент. Класифікація об'єктів виконується в звичайному режимі з використанням стисненого опису, що значно ефективніше за критерієм швидкодії. Імовірність правильної класифікації при цьому зберігається на високому рівні.

## 2.5 Мультимножини у структурному розпізнаванні

Мультимножина – це множина з повторюваними елементами, де один і той же елемент може бути присутнім багаторазово, особливістю мультимножини є поняття кратності входження елемента.

Мультимножиною  $AM$ , визначеною на множині  $A = \{x_1, x_2, \dots\}$  усі елементи  $x_i$ , якого різняться, називається сукупність груп однакових елементів  $AM = \{k_1x_1, k_2x_2, \dots\}, x_iA$ .

Групу однакових елементів  $k_1x_1$ , називають компонентою мультимножини, елементи  $x_i$ , які входять в  $k_1x_1$ , – екземплярами елементів мультимножини. Функція  $k_i$  приймаюча числові значення, визначає число входжень елемента  $x_iA$  в мультимножині  $AM$ . Її також називають функцією кратності або функцією числа примірників мультимножини  $AM$ .



Кажуть, що елемент  $x_i$  належить мультимножині  $AM$  та в мультимножині  $AM$  є рівно  $k$  примірників елемента  $x_i$ , тоді і тільки тоді, коли кратність елемента  $x_i$  дорівнює  $k_i x_i > 0$ .

Коли кратність елемента  $x_i$  дорівнює нулю  $k_i x_i = 0$ , тоді кажуть, що елемент  $x_i$  не міститься в мультимножині  $AM$ . Тим самим приналежність елемента  $x_i$  мультимножині  $AM$  визначається значенням функції кратності.

Основними характеристиками мультимножини є потужність і розмірність. Потужність мультимножини  $AM$  визначається як загальне число примірників всіх його елементів:

$$|AM| = \text{card}AM, \quad (3.1)$$

а розмірність мультимножини  $A$  – як загальне число різних елементів:

$$|AM| = \text{dim}AM. \quad (3.2)$$

Структурний опис об'єкта  $Z = \{z^k\}_{k=1}^s$ ,  $s < \infty$  з універсуму  $AM$ ,  $Z \in AM$ , являє собою кінцеву мультимножину потужності  $s$ , елементи якого – характерні ознаки (ХО) мають структуру:

$$z^k = (x^k, y^k, \lambda^k), \quad (3.3)$$

де  $\lambda \in R^n$ ,  $\lambda = \{\lambda_i\}_{i=1}^n$  – атрибут, який представляє собою числовий вектор розмірності  $n$ , а  $x, y$  – піксельні координати ХО.

Атрибут  $\lambda$  обчислюється таким чином, щоб мати інваріантність до певної групи геометричних перетворень  $G$  [41]. В даний час найбільш популярні методи для формування ХО – SIFT и SURF [41], які формують атрибут, інваріантний до перетворень зсувів, поворотів і масштабу. Класифікація об'єкта на основі його опису  $Z$  із застосуванням структурно-ієрархічного методу (СІМ) розглядається як відображення  $\mathcal{R}: Z \rightarrow Z'$ , де  $Z = \{Z(j)\}$ ,  $\{Z(j)\}_{j=1}^J$  – кінцева множина еталонних описів (база відеоданих), і реалізується рішенням задачі оптимізації:

$$j^* = \arg \max_{j \in \{1, \dots, J\}} \mathfrak{G}[Z, Z(j)], \quad (3.4)$$

де  $\mathfrak{G} \in R^1$  – деяка міра подібності описів.

У відомих підходах визначення подібності  $\mathfrak{G}(Z, Z(j))$  зводиться до обчислення потужності  $\mu(Z_j)$  підмножини  $Z_j \subseteq Z$ , для елементів якого встановлено відповідність множини  $Z(j) \in \square$ , а рішення (3.2) визначається як вибір  $Z_j$  максимальної потужності. Значення частки відданих за клас  $j$  голосів трактується як оцінка апостеріорної ймовірності віднесення  $Z$  до класу  $j$ . Обчислення  $\mu(Z_j)$  в теоретико-множинній моделі пов'язано з голосуванням  $z^k \in Z$  і призводить до побудови на множині  $Z$  представлення  $Z = \bigcup_j Z_j$  у вигляді розбиття ( $Z_i \cap Z_j = \emptyset$ ) або покриття ( $Z_i \cap Z_j \neq \emptyset$ ).

Міри подібності  $\mathfrak{G}(Z, Z(j))$  на основі подання у вигляді множин ХО є більш ефективними в плані завадостійкості в порівнянні з мірами, в основі яких лежить аналіз відповідностей множини або незалежне голосування окремих ХО.

З огляду на те, що значення ХО всередині одного опису можуть бути еквівалентними (повторюватися) через близькість властивостей фрагментів зображень, при аналізі подібності структурних описів доцільним видається використання апарату мультимножин. Застосування мультимножин в прикладних задачах комп'ютерного зору за рахунок обліку повторюваності ХО в описі дозволить забезпечити точність пошуку і розпізнавання зображень в базах візуальних даних.

Апарат мультимножин доцільно використовувати спільно з проведенням процедур попередньої кластеризації або побудови гістограмного уявлення структурного опису [43]. Формування описів у вигляді мультимножин – це різновид побудови системи класів на множині ХО.

При формуванні структурного опису побудова  $U$  як множина різних між собою атомарних структурних елементів є ключовою проблемою, тому що передбачає реалізацію деякої формальної процедури класифікації елементів опису.

В якості одного з природних варіантів можна розглядати базову множину  $U = \{\lambda^i\}$ ,  $\lambda^i \in R^n$  атрибутів ХО у вигляді підмножини  $U \subset R^n$ .

Основою є те, що вектор ХО в точці зображення найбільш точно відображає властивості візуального об'єкта в окремо взятому фрагменті. Класифікація векторів в цьому випадку формалізується у вигляді відображення  $K_\lambda : R^n \rightarrow U$ . У такій моделі  $U$  може бути скінченною або нескінченною множиною і, як правило, формується на основі

конкретного апріорно відомого опису бази  $R$ , в рамках якої здійснюється пошук або розпізнавання об'єктів.

Одним з класичних критеріїв класифікації у векторному просторі є величина норми вектора. Однак дослідження показують, що для класифікації ХО зображень  $i$ , відповідно, для побудови базової множини, норма вектора не є достатньо ефективним критерієм. Системи ХО, отримані технологіями SIFT і SURF, взагалі припускають обробку таким чином, що норми усіх векторів зі сформованого структурного опису приблизно однакові, і тому не можуть виступати критерієм для класифікації.

Іншим способом є уявлення множини атомарних об'єктів, які формують базову множину  $U$  из компонентів вектора ХО. Тут інтегровані описи еталонів з бази  $\square$  сформовані у вигляді гістограм з компонентів вектора ХО, отриманих дискретизацією значень компонентів шляхом розбиття на фіксоване число інтервалів. Обробка даних в цьому підході носить в деякому плані синтетичний характер, тому що передбачає розрізнений аналіз складових єдиного опису, однак метод працездатний і має важливе прикладне значення з точки зору високої швидкодії через істотне скорочення обсягу опису. Перехід до найбільш низького рівня структуризації, при якому руйнуються внутрішні зв'язки між компонентами вектора ХО, виявляється корисним при практичному застосуванні.

Ще одним критерієм, що відображає властивості ХО як елементів візуального опису, є встановлення різновидів (алфавіту, класів) локальних околиць, для яких побудовані ХО: ізольована точка, кут, перетин ліній, т-подібна точка, кінець лінії. [41].

Реалізація такого вельми загального критерію передбачає застосування статистичної обробки значень відгуків відповідних фільтрів, наприклад, SURF. З огляду на величезну різноманітність різновидів локальних околиць зображень  $i$  візуального представлення для одного і того ж типу ХО з перерахованих класів, рішення цього складного завдання навряд чи може бути здійснено однозначно за прийнятний час і з потрібним ступенем достовірності, хоча людський зір справляється з такими завданнями без проблем.

Критерієм, що виключає трудомісткий попередній статистичний аналіз наявних даних з метою якісного навчання системи класифікації, може бути використання величин близькості характеристик аналізованого ХО об'єкта і кластерів ХО, сформованих для еталонних даних відеобаз. Цей критерій відповідає принципу ближнього сусіда. Класи базових елементів мультимножини сформуємо для реалізації такого аналізу на основі кластеризації множини структурних елементів бази еталонів, що забезпечує погодження на конкретний предмет аналізу і забезпечує необхідну ефективність в плані швидкодії і ймовірності правильного розпізнавання. Реалізація кластеризації у вигляді непересічних

підмножин (розбиття) визначає між елементами вихідної множини певне відношення еквівалентності. Об'єкти одного кластера (класу) розглядаються як однакові, а представник кластера (атомарний об'єкт) природним чином представляється як елемент базової множини  $U$ .

Кластеризація може бути виконана як на основі векторів ХО, так і на базі їх характеристик, включаючи гістограмне уявлення.

Порівняння двох мультимножин як векторів відбувається за манхеттенською метрикою:

$$d(p, q) = |p - q| = \sum_{i=1}^n |p_i - q_i|. \quad (3.5)$$

## 3 РЕЗУЛЬТАТИ ЕКСПЕРИМЕНТАЛЬНИХ ДОСЛІДЖЕНЬ

### 3.1 Обґрунтування вибору середовища програмної реалізації

У рамках магістерської роботи розроблений метод для порівнювання та обчислення релевантності зображень. Для реалізації було обране середовище PyCharm Community Edition 2017.2.3. Це обумовлено тим, що в PyCharm зручно писати код на мові Python.

Можливості PyCharm:

- статичний аналіз коду, підсвічування синтаксису і помилок;
- навігація по проекту і вихідного коду: відображення файлової структури проекту, швидкий перехід між файлами, класами, методами і використаннями методів;
- рефакторинг: перейменування, вилучення методу, введення змінної, введення константи, підйом і спуск методу і таке інше;
- інструменти для веб-розробки з використанням фреймворку Django;
- вбудований відладчик для Python;
- вбудовані інструменти для юніт-тестування;
- розробка з використанням Google App Engine;
- підтримка систем контролю версій: загальний користувальницький інтерфейс для Mercurial, Git, Subversion, Perforce і CVS з підтримкою списків змін і злиття.

Python – це один з найбільш популярних сучасних мов програмування. Він придатний для вирішення різноманітних завдань і пропонує ті ж можливості, що і інші мови програмування: динамічність, підтримку ООП і крос-платформенність.

Завдяки легкому для читання синтаксису, код читається легше, що означає, багаторазове його використання і обслуговування виконується набагато простіше, ніж використання програмного коду на інших мовах сценаріїв. Python містить найсучасніші механізми багаторазового використання програмного коду, яким є ООП.

Так само Python підтримує система управління пакетами (pip), яка використовується для установки і управління програмними пакетами, які написані на Python.

Більшість дистрибутивів Python вже містять pip. Якщо, pip відсутня, то його можна встановити за допомогою системи управління пакетами або через cURL, утиліту для завантаження через інтернет.

Одне з головних переваг pip – це простота інтерфейсу командного рядка, яка дозволяє встановити пакети Python простою командою: Так само просто і видаляти пакет.

Головне, що `pip` надає можливість управляти всіма пакетами і їх версіями. Це зазвичай робиться за допомогою файлу `requirements.txt`. Що дозволяє ефективно відтворювати весь необхідний список пакетів в окремому оточенні (наприклад, на іншому комп'ютері) або в віртуальному оточенні. Це досягається за допомогою правильно відформатованого файлу `requirements.txt` і команди. Установка деяких пакетів для конкретних версій `python`, де `{version}` замінюється на 2, 3, 3.6, і таке інше.

Переваги `pip`:

- завантаження всіх необхідних пакетів перед їх установкою;
- забезпечення зворотного зв'язку;
- ведення історії дій, що робляться;
- висновок повідомлень про помилки (за традицією `Python`).

Інструмент `pip` дійсно простий у використанні і здатний усунути багато актуальних проблем і спростити виконання багатьох завдань, з якими регулярно доводиться стикатися розробникам додатків.

`PyCharm` забезпечує зручне використання цієї можливості у терміналі.

Бібліотеку `OpenCV` неможливо завантажити подібним способом, проте за допомогою `pip` можна контролювати її версію. Так чи інакше ця бібліотека представлена широким функціоналом роботи з `Python`. Так само варто відзначити доступну документацію, яка робить розробку набагато швидше і зручніше.

## 3.2 Результати експериментальних досліджень

Використання апарату мультимножин забезпечує необхідний рівень достовірності класифікації при порівнянні зображень.

У проведеному комп'ютерному експерименті мультимножини використовувалися на двох рівнях опису.

На першому рівні кожна ХО з опису представляється у вигляді гістограми значень компонент. Базова множина формувалася на основі 17-ти діапазонів значень компонент детектора SURF, обчислених для повної бази відеоінформації. Таким чином було здійснено перехід від ХО до гістограмної ознаки (ГО). Експерименти здійснювалися для бази відеоданих акваріумних риб. Досліджувана база містить 10 зображень риб. Дослідження показали, що застосування модифікації з переходом в простір ГО скорочує середній час класифікації зображення приблизно в 12 разів у порівнянні з традиційним методом.

Другий рівень обробки включає розрахунок норм гістограм для ГО, отриманих на першому рівні опису. Базова множина побудована для кожного з еталонних описів у вигляді мультимножини. В результаті кожен з еталонів був представлений гістограмою другого рівня.

Запропонована дворівнева система перетворення ХО, як показали експерименти, скорочує час класифікації більш ніж в 8 разів у порівнянні з модифікацією і приблизно в 90 разів в порівнянні з традиційним методом SURF. При цьому обсяг опису скорочується в сотні разів (в середньому в 450) в порівнянні зі стандартним класичним уявленням.

При низькому рівні перешкод всі розглянуті модифікації забезпечують безпомилкову класифікацію для розглянутих зображень. Експериментально оцінювалася також перешкодозахищеність традиційного методу ORB.

Комп'ютерне моделювання показало, що для досліджуваної бази до рівня ймовірності  $\gamma \leq 0,9$  ймовірність  $P$  правильної класифікації для традиційного детектора SURF становить  $P = 1$ , тобто навіть по 10% опису здійснюється правильна класифікація.

Зауважимо, що адитивні спотворення в даному експерименті відсутні. Такий же рівень ймовірності класифікації  $P = 1$  досягається і для модифікованого одноступінчастого методу із застосуванням ГО. У той же час для двухфазного методу ймовірність  $P = 1$  досягається лише при ймовірності зникнення  $\gamma \leq 0,4$ , значення  $P = 0,9$  при  $\gamma = 0,5$ .

Таким чином, маємо, що метод, який використовує ГО другого рівня, здатний правильно класифікувати вхідне зображення на основі 60% неспотвореного опису, а далі його ефективність знижується. Однак і рівень 60% спотворень є досить прийнятним на практиці. У той же час з точки зору критерію часу обробки цей метод найефективніший.

Міри подібності на основі подання у вигляді множин ХО є більш ефективними в плані завадостійкості в порівнянні з заходами, в основі яких лежить аналіз відповідностей множини або незалежне голосування окремих ХО. З огляду на те, що значення ХО всередині одного опису можуть бути еквівалентними через схожість властивостей фрагментів зображень, при аналізі подібності структурних описів доцільним видається використання апарату мультимножин. Застосування мультимножин в прикладних задачах комп'ютерного зору за рахунок обліку повторюваності ХО в описі дозволить забезпечити точність пошуку і розпізнавання зображень в базах відеоданих.

Експерименти проводилися для бази зображень домашніх риб. Попередня обробка бази полягала в наступному.

Було отримано масив ключових точок зображень риб за допомогою метода ORB (Oriented FAST and Rotated BRIEF), який реалізовано у модулі проекту «MyTransform.py» (ДОДАТОК А).

Також як і мультимножини можна застосовувати функції Уолша (ДОДАТОК Б) та отримати значення, за допомогою яких можна отримати вектори з ХО та порівняти їх.

Сформовано та вибрано 5 функцій Уолша:

- а) усі 1;
- б) сто двадцять вісім 1 та  $-1$ ;
- в) по шістдесят чотири 1 та  $-1$ ;
- г) по чотири 1 та  $-1$ ;
- г) 1 та  $-1$  по черзі.

Обчислюємо скалярний добуток усіх масивів ключових точок з функціями Уолша а)- г).

У результаті отримано масив з результатами (для кожної точки своє значення функції Уолша).

За манхеттенською метрикою вираховуємо відстань, з якої потрібно знайти мінімуми.

З цих мінімумів робимо висновки щодо схожості зображень обираючи поріг значення за 76. Якщо 90% значень менше цього порогу, то риби схожі між собою.

Перевагами цього алгоритму можна виділити: вичислення відбувається з цілими значеннями, значне поліпшення зрівняння зображень за допомогою використання векторів.

Серед недоліків можна виділити узагальнення представлень через інтегральне представлення, що виключає врахування дрібниць.

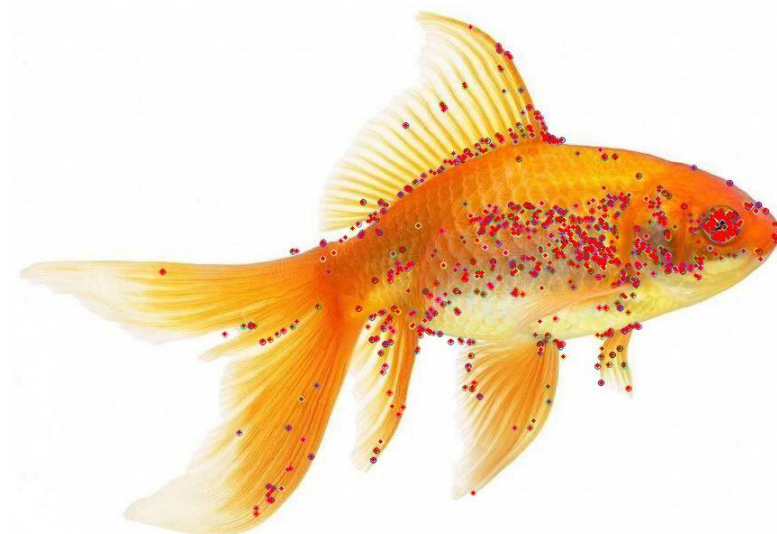


Рисунок 4.1 – Зображення для дослідження №1



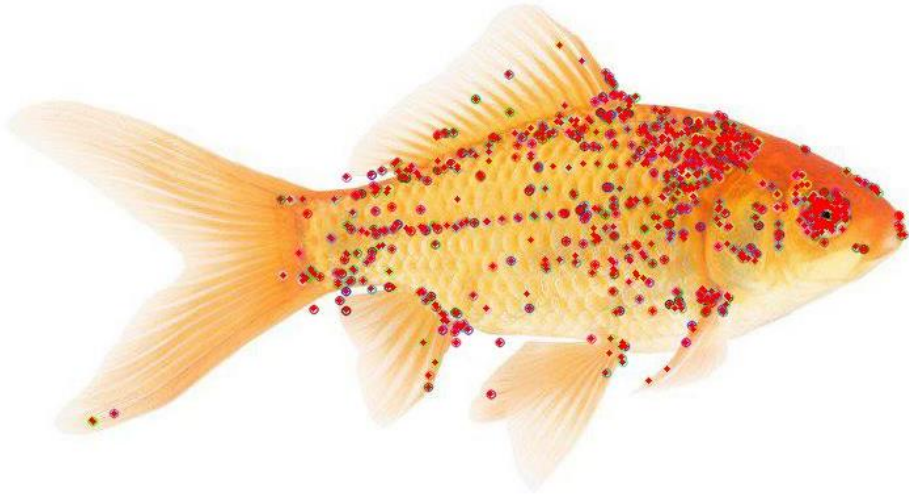


Рисунок 4.2 – Зображення для дослідження №2



Рисунок 4.3 – Зображення для дослідження №3



Рисунок 4.4 – Зображення для дослідження №4

Провівши порівняльний аналіз методу, який є предметом дослідження магістерської роботи, з іншими методами порівняння зображень, було зроблено висновок, що даний варіант розробки є конкурентоспроможним. Точність розпізнавання досягає 80% з усіх досліджуваних варіантів, а час порівняння двох зображень риб становить менше 0,32 секунди.

## 4 ОХОРОНА ПРАЦІ

В даному розділі проведено аналіз потенційних небезпечних та шкідливих виробничих факторів, причин пожеж. Розглянуті заходи, які дозволяють забезпечити гігієну праці і виробничу санітарію. На підставі аналізу розроблені заходи з техніки безпеки та рекомендації з пожежної профілактики.

### 4.1 Аналіз небезпечних та шкідливих факторів при виробництві (експлуатації) виробу

На підставі аналізу небезпечних та шкідливих факторів при виробництві (експлуатації), пожежної безпеки можуть бути надалі вирішені питання необхідності забезпечення працюючих достатньою кількістю освітлення, вентиляції повітря, організації заземлення, тощо.

Роботу, пов'язану з ЕОП з ВДТ, у тому числі на тих, які мають робочі місця, обладнані ЕОМ з ВДТ і ПП, виконують із забезпеченням виконання ДСанПіН 3.3.2.007-98 [41], які встановлюють вимоги безпеки до обладнання робочих місць, до роботи із застосуванням ЕОМ з ВДТ і ПП. Переважно роботи за проектами виконують у кабінетах чи інших приміщеннях, де використовують різноманітне елект-рообладнання, зокрема персональні комп'ютери (ПК) та периферійні пристрої.

Основними робочими характеристиками персонального комп'ютера є наступні:

- робоча напруга  $U = +220\text{В} \pm 5\%$ ;
- робочий струм  $I = 2\text{А}$ ;
- споживана потужність  $P = 350\text{Вт}$ .

Робочі місця мають відповідати вимогам Державних санітарних правил і норм роботи з ві-зуальними дисплейними терміналами електронно-обчислювальних машин, затверджених постановою Головного державного санітарного лікаря України від 10.12.98 N 7 [41].

За умов роботи з ПК виникають наступні небезпечні та шкідливі чинники: несприятливі мі-крокліматичні умови, освітлення, електромагнітні випромінювання, забруднення повітря шкідливими речовинами (джерелом, яких можуть бути: принтер, сканер та інші джерела виділення багатьох хімічних речовин - напр., озону, оксидів азоту

та аерозолів високодисперсних частинок тонера), шум, вібрація, електричний струм, електростатичне поле, напруженість трудового процесу та інше.

Аналіз небезпечних та шкідливих виробничих факторів виконується у табличній формі (табл. 4.1).

Таблиця 4.1 – Аналіз небезпечних і шкідливих виробничих факторів

Небезпечні і шкідливі виробничі фактори	Джерела факторів (види робіт)	Кількіс на оцінка	Нормативні документи
1	2	3	4
<b>фізичні</b>			
- підвищена температура поверхонь обладнання	експлуатація ЕОМ, принтерів, сканерів чи/або серверного обладнання для роботи	2	[42]
- підвищений рівень шуму на робочому місці	-//-	2	[43]
- підвищена або знижена рухливість повітря	-//-	1	[42]
- підвищений рівень напруги електричної мережі, замикання якої може відбутися через тіло людини	-//-	4	[45]
- підвищена напруженість електричного поля	-//-	2	[45]
- недостатність природного світла	порушення умов праці (вимог до приміщень)	2	[46]
- недостатнє освітлення робочої зони	порушення гігієнічних параметрів виробничого середовища	3	[46]
<b>психофізіологічні:</b>			
- нервово-психічна перевантаження (розумове, перенапруження аналізаторів-зорових)	- пошук інформації для постановки теми; - пошук та аналіз аналогів і літератури; - пошук наявних технологій, моделювання та аналіз алгоритмів; - виконання роботи за темою диплома, тестування; - оформлення роботи	4	[41] [50]
- фізичні (статичне – сидіння)	порушення умов праці (організації місця праці- сидіння користувача, ) та організації робочого часу - безпервна робота)	2	[41] [50]

## 4.2 Гігієнічні вимоги до параметрів виробничого середовища

### 4.2.1 Мікроклімат

Оптимальні значення для температури, відносної вологості й рухливості повітря для зазначеного робочого місця відповідають ДСН 3.3.6.042-99 [42] і наведені в табл. 4.2:

Таблиця 4.2 – Норми мікроклімату робочої зони об'єкту

Період року	Категорія робіт	Температура С <sup>0</sup>	Відносна вологість %	Швидкість руху повітря, м/с
Холодна	легка-1 а	22 - 24	40 – 60	0,1
Тепла	легка-1 а	23 - 25	40 – 60	0,1

Дане приміщення обладнане системами опалення, кондиціонування повітря або припливно-витяжною вентиляцією. У приміщенні на робочому місці забезпечуються оптимальні значення параметрів мікроклімату: температури, відносної вологості й рухливості повітря у відповідності до ДСН 3.3.6.042-99 [42]. Для забезпечення оптимальних параметрів мікроклімату в приміщенні проводяться перерви в роботі співробітників, з метою його провітрювання. Існують спеціальні системи кондиціонування, які забезпечують підтримання в приміщенні балансу оптимальних параметрів мікроклімату. Контроль параметрів мікроклімату в холодний і теплий період року здійснюється не менше 3-х разів на зміну (на початку, середині, в кінці).

### 4.2.2 Освітлення

У приміщенні, де розташовані ЕОМ передбачається природне бічне освітлення, рівень якого відповідає ДБН В. 2.5-28:2018 [46]. Джерелом природного освітлення є сонячне світло. Регулярно повинен проводитися контроль освітленості, який підтверджує, що рівень освітленості задовольняє ДБН і для даного приміщення в світлий час доби достатньо природного освітлення.

*Розрахунок освітлення.*

Для будівель виробництв світловий коефіцієнт приймається в межах 1/6 - 1/10:

$$\sqrt{a^2 + b^2} \cdot S_b = (1/8 \div 1/10) \cdot S_n \quad (4.1)$$

де  $S_b$  – площа віконних прорізів, м<sup>2</sup>;

$S_n$  – площа підлоги, м<sup>2</sup>.

$$S_n = a \cdot b = 5 \cdot 5 = 25 \text{ м}^2$$

$$S_{\text{вік}} = 1/8 \cdot 25 = 3,125 \text{ м}^2$$

Приймаємо 2 вікна площею  $S = 1,6 \text{ м}^2$  кожне.

Світильники загального освітлення розташовуються над робочими поверхнями в рівномірно-прямокутному порядку. Для організації освітлення в темний час доби передбачається обладнати приміщення, довжина якого складає 5 м, ширина 5 м, світильниками ЛПО2П, оснащеними лампами типа ЛБ (дві по 80 Вт) з світловим потоком 5400 лм кожна.

Розрахунок штучного освітлення виробляється по коефіцієнтах використання світлового потоку, яким визначається потік, необхідний для створення заданої освітленості при загальному рівномірному освітленні. Розрахунок кількості світильників  $n$  виробляється по формулі (4.2):

$$n = \frac{E \cdot S \cdot Z \cdot K}{F \cdot U \cdot M} \quad (4.2)$$

де  $E$  – нормована освітленість робочої поверхні, визначається нормами – 300 лк;

$S$  – освітлювана площа, м<sup>2</sup>;  $S = 25 \text{ м}^2$ ;

$Z$  – поправочний коефіцієнт світильника ( $Z = 1,15$  для ламп розжарювання та ДРЛ;  $Z = 1,1$  для люмінесцентних ламп) приймаємо рівним 1,1;

$K$  – коефіцієнт запасу, що враховує зниження освітленості в процесі експлуатації – 1,5;

$U$  – коефіцієнт використання, залежний від типу світильника, показника індексу приміщення і т.п. – 0,575

$M$  – число люмінесцентних ламп в світильнику – 2;

$F$  – світловий потік лампи – 5400лм (для ЛБ-80).

Підставивши числові значення у формулу (4.2), отримуємо:

$$n = \frac{300 \cdot 25 \cdot 1,1 \cdot 1,5}{5400 \cdot 0,575 \cdot 2} \approx 2.$$

Приймаємо освітлювальну установку, яка складається з 2-х світильників, які складаються з двох люмінесцентних ламп загальною потужністю 160 Вт, напругою – 220 В.

#### 4.2.3 Шум та вібрація, електромагнітне випромінювання

Рівень шуму, що супроводжує роботу користувачів персональних комп'ютерів (зумовлений як роботою системних блоків, клавіатури, так і друкуванням на принтерах, а також зовнішніми чинниками), коливається у межах 50–65 дБА [43]. Шум такої інтенсивності на тлі високого ступеня напруженості праці негативно впливає на функціональний стан користувачів. Тому на практиці рекомендують знижувати фактичний рівень шуму у приміщеннях, де створюють комп'ютерні програми, виконують теоретичні та творчі роботи, проводять навчання до 40 дБА, а в приміщеннях, де виконують роботу, що потребує зосередженості, — до 55 дБА. У залах опрацювання інформації та комп'ютерного набору рівні шуму не повинні перевищувати 65 дБА.

Шум часто є причиною зниження рівня працездатності, підвищення рівня загальної та професійної захворюваності, частоти виробничих травм. Шум є загальнобіологічним подразником, який негативно впливає на всі органи і системи організму. У разі тривалого систематичного впливу шуму може виникнути патологія з переважним ураженням слуху, центральної нервової і серцево-судинної систем.

Для зниження шуму на шляху його поширення передбачається розміщення в приміщенні штучних поглиначів. Для зниження рівня шуму стелю або стіни вище 1.5 - 1.7 метра від підлоги повинні облицьовуватися звукопоглинальним матеріалом з максимальним коефіцієнтом звукопоглинання в області частот 63-8000 Гц. Додатковим звукопоглинанням в КВТ можуть бути фіранки, підвішені в складку на відстані 15-20 см. Від огорожі, виконані з щільної, важкої тканини. У приміщенні з ЕОМ коректований рівень звукової потужності не перевищує 45 дБА. Оскільки рівень шуму не перевищує гранично допустимих величин, які встановлені санітарними нормами, заходи для зниження шуму не проводяться.

Віброізоляцію можливо здійснювати за допомогою спеціальної прокладки під системний блок, який послаблює передачу вібрацій робочого столу. Вібрація на робочому

місці в приміщенні, що розглядається, відповідає нормам [43]. Допустимий рівень вібрацій на робочому місці: для 1 ступеня шкідливості до 3 дБ; для 2-3 - 1-6 дБ; для 3 - більше 6 дБ.

Для захисту від електромагнітного випромінювання передбачаються наступні заходи:

- 1) застосування нових плазмових моніторів, LG W2271TC,
- 2) віддалення робочого місця не менше, ніж на 0,4-0,5 м, оскільки напруженість електричного поля зменшується при віддаленні від джерела поля,
- 3) встановлення раціональних режимів роботи персоналу (обмеження часу перебування),
- 4) раціональне розміщення в робочому приміщенні устаткування, що випромінює електромагнітну енергію.

#### 4.2.4 Вентилювання

У приміщенні, де знаходяться ЕОМ, повітрообмін реалізується за допомогою природної організованої вентиляції (вентиляційні шахти) і установки в віконному отворі автономного кондиціонера БК-2000. Цей метод забезпечує приток потрібної кількості свіжого повітря, що визначається в СНіП (30 м<sup>3</sup> на годину на одного працюючого).

Також має здійснюватися провітрювання приміщення, в залежності від погодних умов, тривалість повинна бути не менше 10 хв. Найкращий обмін повітря здійснюється при наскрізному провітрюванні.

### 4.3 Заходи з організації виробничого середовища та попередження виникнення надзвичайних ситуацій

Розрахунок захисного заземлення (забезпечення електробезпеки будівлі).

Загальний опір захисного заземлення визначається за формулою:

$$R_{3zn} = \frac{R_z \cdot R_n}{R_n \cdot n \cdot \eta_z + R_z \cdot \eta_n}, \quad (4.3)$$



де  $R_3$  - опір заземлення, якими когут бать труби, опори, кути і т.п., Ом;

$R_{ш}$  - опір опори, яке з'єднує заземлювачі, Ом;

$n$  - кількість заземлювачів;

$\eta_3$  - коефіцієнт екранування заземлювача; приймається в межах  $0,2 \div 0,9$ ;  $\eta_3 = 0,7$

$\eta_{ш}$  - коефіцієнт екранування сполучної стійки; приймається в межах  $0,1 \div 0,7$ ;  $\eta_{ш} = 0,5$ ;

Опір заземлення визначається за формулою:

$$R_3 = \frac{\rho}{2\pi \cdot l} \cdot \left( \ln \frac{2 \cdot l}{d} + \frac{1}{2} \ln \frac{4 \cdot t + l}{4 \cdot t - l} \right), \quad (4.4)$$

де  $\rho$  - питомий опір ґрунту, залежить від типу ґрунту, Ом·м;

для піску -  $400 \div 700$  Ом·м; приймаємо  $\rho = 400$  Ом·м;

$l$  - довжина заземлювача, м; для труб - 2-3 м;  $l = 3$  м;

$d$  - діаметр заземлювача, м; для труб - 0,03-0,05 м;  $d = 0,05$  м;

$t$  - відстань від середини забитого в ґрунт заземлювача до рівня землі, м;

$t = 2$  м.

$$R_3 = \frac{400}{2 \cdot 3,14 \cdot 3} \left( \ln \frac{2 \cdot 3}{0,05} + \frac{1}{2} \ln \frac{4 \cdot 2 + 3}{4 \cdot 2 - 3} \right) = 110, \text{ Ом}$$

Опір смуги, що з'єднує заземлювачі, визначається за формулою:

$$R_{ш} = \frac{\rho}{2\pi \cdot L} \cdot \ln \frac{2 \cdot L^2}{b \cdot t_1}, \quad (4.5)$$

де  $L$  - довжина смуги, що з'єднує заземлювачі (м) і приблизно дорівнює периметру будівлі:  $P_{буд} = 42 \cdot 2 + 38 \cdot 2 = 160$  м;  $L = 160$  м;

$b$  - ширина смуги, м;  $b = 0,03$  м;

$t_1$  - глибина заземлення від рівня землі, м;  $t_1 = 0,5$  м.

$$R_{ш} = \frac{400}{2 \cdot 3,14 \cdot 160} \cdot \ln \frac{2 \cdot 160^2}{0,03 \cdot 0,5} = 5,99, \text{ Ом}$$

Кількість заземлювачів захисного заземлення визначається за формулою:

$$n = \frac{2 \cdot R_3}{4 \cdot \eta_3}, \quad (4.6)$$

де 4 - допустимий загальний опір, Ом;

2 - коефіцієнт сезонності.

Визначаємо загальний опір захисного заземлення:

$$R_{ззп} = \frac{110 \cdot 5,99}{5,99 \cdot 79 \cdot 0,7 + 110 \cdot 0,5} = 1,7 \text{ Ом}$$

Висновок: дане захисне заземлення буде забезпечувати електробезпеку будівлі, так як виконується умова:  $R_{ззп} < 4 \text{ Ом}$ .

3) При виникненню пожеж при роботі на ПЕОМ від таких можливими джерел запалювання як:

- іскри і дуги коротких замикань;
- перегрів провідників, резисторів та інших радіодеталей ПЕОМ, від тривалої перевантаження та наявність перехідного опору;
- іскри при розмиканні і розмиканні ланцюгів;
- розряди статичної електрики;
- необережному поводженню з вогнем, а також вибухи газо-повітряних і пароповітряних сумішей.

Важливу увагу слід звернути на пожежну безпеку підприємства в цілому і окремих його приміщень. В приміщеннях не повинно накопичуватися сміття, непотрібний папір, мотлох та ін. речі, які не використовуються у виробничому процесі. Наявний вільний аварійний вихід за межі приміщення в разі пожежі, бути передбачені вогнегасники. Вони повинні бути в робочому стані і перевірятися згідно з нормами. У приміщеннях повинна бути пожежна сигналізація, вогнегасник. У разі виникнення пожежі необхідно повідомити в найближчу пожежну частину, убезпечити інших працівників і по можливості прийняти кроки по запобіганню можливих наслідків та усуненню пожежі.

### 4.3 Охорона навколишнього природного середовища

Діяльність за темою магістерської роботи, а саме: оптимізація запитів до бази даних в процесі її виконання впливає на навколишнє природне середовище і регламентується нормами діючого законодавства: Законом України «Про охорону навколишнього природного середовища» [52], Законом України «Про забезпечення санітарного та епідемічного благополуччя населення» [53], Законом України «Про відходи» [54].

В процесі роботи виникають процеси поводження з відходами ІТ галузі. Нижче надано перелік відходів, що утворюються в процесі роботи:

- відпрацьовані люмінесцентні лампи - I клас небезпеки;
- змінні носії інформації - IV клас небезпеки;
- відходи друкуючих пристроїв - IV клас небезпеки;
- макулатура - IV клас небезпеки;
- побутові відходи - IV клас небезпеки.

## ВИСНОВКИ

В магістерській роботі була досліджена актуальна задача розпізнавання зображень на основі застосування розроблених методів і моделей трансформації ОО з метою побудови ефективних мір подібності.

При проведенні досліджень отримані наступні результати:

1. Встановлено, що аналіз застосовуваних ОО і виявлення особливостей бази зображень поряд з застосовуваним методом значно впливають на результат розпізнавання. Застосування процедур перетворення опису або переходу в новий простір ознак сприяє поліпшенню характеристик розпізнавання;

2. Обраний метод побудови описів на основі функцій Уолша дозволяє проводити розпізнавання об'єктів з високою швидкістю без зниження перешкодозахищеності;

3. Розроблена модель стиснення ОО за результатами експерименту дозволяє зберегти безпомилкове розпізнавання на рівні 40% від усього числа точок опису. Це допускає рівень стиснення вихідного простору ознак опису в 2,5 рази;

4. Удосконалена модель побудови мір подібності на основі принципу найближчих сусідів забезпечує високу ймовірність правильної класифікації. Модель подібності описів забезпечує порівнянну з традиційними підходами перешкодозахищеність;

5. Реалізація запропонованих моделей і методів дозволяє вирішити ряд прикладних задач. Досягнуто необхідний рівень достовірності та швидкодії при класифікації зображень риб.

В результаті проведеної роботи було зроблено аналіз умов праці, шкідливих та небезпечних чинників, з якими стикається робітник. Було визначено параметри і певні характеристики приміщення для роботи над запропонованим проектом написаному в дипломній роботі, описано, які заходи потрібно зробити для того, щоб дане приміщення відповідало необхідним нормам і було комфортним і безпечним для робітника. Приведені рекомендації щодо організації робочого місця, а також важливу інформацію щодо пожежної та електробезпеки, охорони навколишнього природного середовища. Було наведено значення температури, вологості й рухливості повітря, необхідна кількість і потужність ламп та інші параметри, значення яких впливає на умови праці робітника, а також – наведені інструкції з охорони праці, техніки безпеки при роботі на комп'ютері.

## ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАНЬ

- 1) Шапиро, Л., Стокман, Д., Богуславский, А. А., & Соколов, С. М. (2013). Компьютерное зрение. БИНОМ. Лаборатория знаний.
- 2) Форсайт, Д., & Понс, Ж. (2004). Компьютерное зрение. Современный подход (р. 928). М.: ИД Вильямс.
- 3) Машталир, В. П. (2001). Точечно-множественные методы обработки информации. Х.: Бизнес Информ.
- 4) Вовк С.М., Гнатушенко В.В., Бондаренко М.В. Методи обробки зображень та компютерний зір : навч. посіб. / С.М. Вовк, В.В. Гнатушенко, М.В. Бондаренко. – Д. : ЛПРА, 2016. – 148 с.
- 5) Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. IEEE transactions on information theory, 13(1), 21-27.
- 6) Гонсалес, Р., & Вудс, Р. (2019). Цифровая обработка изображений. Litres.
- 7) Павлидис, Т. (1986). Алгоритмы машинной графики и обработки изображений. Радио и связь.
- 8) Шлезингер, М., & Главач, В. (2004). Десять лекций по статистическому и структурному распознаванию. Киев: Наукова думка, 546.
- 9) Jensen, J. R., & Lulla, K. (1987). Introductory digital image processing: a remote sensing perspective.
- 10) Загоруйко, Н. Г. (1999). Прикладные методы анализа данных и знаний. Новосибирск: Изд-во Ин-та математики.
- 11) Гороховатский, В. А., Власенко, Н. В., & Полякова, Т. В. (2014). Построение модификаций и анализ свойств структурных описаний при распознавании изображений. Вестник Национального технического университета Харьковский политехнический институт. Серия: Информатика и моделирование, (35 (1078)).
- 12) Пупков, К. А., Егупов, Н. Д., Пупков, К. А., Егупов, Н. Д., Баркин, А. И., Зайцев, А. В., ... & Краснощеченко, В. И. (2004). Методы классической и современной теории автоматического управления. Московский государственный технический университет имени НЭ Баумана (национальный исследовательский университет).
- 13) Ерьсько, Ю. Н. (2002). Локализация изображений в автоматических визирах. М.: Компания Спутник+.
- 14) Браверман, Э. М., & Мучник, И. Б. (1983). Структурные методы обработки эмпирических данных. Наука. Гл. ред. физ.-мат. лит..

- 15) Ширман, Я. Д., Горшков, С. А., Лещенко, С. П., Братченко, Г. Д., & Орленко, В. М. (1996). Методы радиолокационного распознавания и их моделирование. *Зарубежная радиоэлектроника*, (11), 3.
- 16) Pratt W.K. 2016. *Digital Images Processing*. Third edition. Wiley.
- 17) Баклицкий, В. К., Бочкарев, А. М., & Мусьяков, М. П. (1986). Методы фильтрации сигналов в корреляционно-экстремальных системах навигации. М.: Радио и связь, 216.
- 18) Айвазян, С. А., Бухштабер, В. М., Енюков, И. С., & Мешалкин, Л. Д. (1989). *Прикладная статистика: Классификация и снижение размерности*.
- 19) Delmerico, J. A., David, P., & Corso, J. J. (2011). Building facade detection, segmentation, and parameter estimation for mobile robot localization and guidance. 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems.
- 20) Воронцов, К. В. Лекции по методам отбора и синтеза признаков и моделей.
- 21) Lowe, D. G. (1999, September). Object recognition from local scale-invariant features. In *iccv* (Vol. 99, No. 2, pp. 1150-1157).
- 22) Mikolajczyk, K., & Schmid, C. (2005). A performance evaluation of local descriptors.
- 23) Kim, S., & Kweon, I. S. (2006, January). Biologically motivated perceptual feature: Generalized robust invariant feature. In *Asian Conference on Computer Vision* (pp. 305-314). Springer, Berlin, Heidelberg.
- 24) Harris, C. G., & Stephens, M. (1988, August). A combined corner and edge detector. In *Alvey vision conference* (Vol. 15, No. 50, pp. 10-5244).
- 25) Schmid, C., & Mohr, R. (1997). Local grayvalue invariants for image retrieval. *IEEE transactions on pattern analysis and machine intelligence*, 19(5), 530-535.
- 26) Azimi, S., Lall, B., & Gandhi, T. K. (2019, May). Performance Evaluation of 3D Keypoint Detectors and Descriptors for Plants Health Classification. In *2019 16th International Conference on Machine Vision Applications (MVA)* (pp. 1-6). IEEE.
- 27) Koenderink, J. J., & van Doorn, A. J. (1987). Representation of local geometry in the visual system. *Biological cybernetics*, 55(6), 367-375.
- 28) Köthe, U. (2003, September). Edge and junction detection with an improved structure tensor. In *Joint Pattern Recognition Symposium* (pp. 25-32). Springer, Berlin, Heidelberg.
- 29) Wallraven, C., Caputo, B., & Graf, A. (2003, October). Recognition with local features: the kernel recipe. In *null* (p. 257). IEEE.

- 30) Gorokhovatskiy, A. V., Gorokhovatskiy, V. A., Vlasenko, A. N., & Vlasenko, N. V. (2014). Quality Criteria for Multidimensional Object Recognition Based Upon Distance Matrices. *Telecommunications and Radio Engineering*, 73(18).
- 31) Barghout, L., & Sheynin, J. (2013). Real-world scene perception and perceptual organization: Lessons from Computer Vision. *Journal of Vision*, 13(9), 709.
- 32) Федотов Н.Г. Методы стохастической геометрии в распознавании образов.- М.:Радио и связь, 1990.- 144 с.
- 33) Szeliski, R. (2010). *Computer vision: algorithms and applications*. Springer Science & Business Media.
- 34) Даревич, Р. Р. (2005). Підвищення ефективності інтелектуального аналізу тексту шляхом зважування понять в моделі онтології. *Искусственный интеллект*, (3), 571-577.
- 35) Belongie, S., Malik, J., & Puzicha, J. (2002). Shape matching and object recognition using shape contexts. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (4), 509-522.
- 36) Грузман, И. С., Киричук, В. С., Косых, В. П., Перетягин, Г. И., & Спектор, А. А. (2000). *Цифровая обработка изображений в информационных системах*. Новосибирск: Изд-во НГТУ.
- 37) Gorokhovatskiy, V. A. (2016). Efficient Estimation of Visual Object Relevance during Recognition through their Vector Descriptions. *Telecommunications and Radio Engineering*, 75(14).
- 38) Gorokhovatsky, V. (2014). *Structural analysis and intellectual data processing in computer vision*. SMIT: Kharkiv, Ukraine.
- 39) Lutz, M. (2013). *Learning python: Powerful object-oriented programming*. " O'Reilly Media, Inc."
- 40) Zelle, J. M. (2004). *Python programming: an introduction to computer science*. Franklin, Beedle & Associates, Inc.
- 41) Державні санітарні норми і правила. ДСанПіН 3.3.2.007-98 «Гігієнічні вимоги до організації роботи з візуальними дисплейними терміналами електронно-обчислювальних машин» Режим доступу: WWW. URL: <https://zakon.rada.gov.ua/rada/show/v0007282-98>
- 42) Державні санітарні норми України. ДСН 3.3.6.042-99 «Санітарні норми мікроклімату виробничих приміщень» Режим доступу: WWW. URL: <https://zakon.rada.gov.ua/rada/show/va042282-99>

43) Державні санітарні норми України. ДСН 3.3.6.037-99 «Санітарні норми виробничого шуму, ультразвуку та інфразвуку» Режим доступу: WWW. URL: <https://zakon.rada.gov.ua/rada/show/va037282-99>

44) Державні санітарні норми України. ДСН 3.3.6.039-99 «Санітарні норми виробничої загальної та локальної вібрації» Режим доступу: WWW. URL: <https://zakon.rada.gov.ua/rada/show/va039282-99>

45) Державний стандарт України. ГОСТ 13109-97 «Электрическая энергия. Совместимость технических средств электромагнитных. Нормы качества электроэнергоснабжения общего назначения» Режим доступу: WWW. URL: [http://odz.gov.ua/lean\\_pro/standardization/files/elektromagnitnaja\\_sovmestimost\\_2014\\_03\\_11\\_1.pdf](http://odz.gov.ua/lean_pro/standardization/files/elektromagnitnaja_sovmestimost_2014_03_11_1.pdf)

46) Державні будівельні норми України. ДБН В.2.5-28:2018 «Природне і штучне освітлення» Режим доступу: WWW. URL: [https://okna.ua/img\\_all/okna.ua/dbn-V-2-5-28-2018-ed.pdf](https://okna.ua/img_all/okna.ua/dbn-V-2-5-28-2018-ed.pdf)

47) Нормативно-правові акти з охорони праці. НПАОП 40.1-1.21-98 «Правила безпечної експлуатації електроустановок споживачів» Режим доступу: WWW. URL: <https://zakon.rada.gov.ua/laws/show/z0093-98>

48) Державні будівельні норми України. ДБН В.2.5-67:2013 «Опалення, вентиляція та кондиціонування» Режим доступу: WWW. URL: [https://dnaop.com/html/32609/doc-%D0%94%D0%91%D0%9D\\_%D0%92.2.5-67\\_2013](https://dnaop.com/html/32609/doc-%D0%94%D0%91%D0%9D_%D0%92.2.5-67_2013)

49) Державний стандарт України. ГОСТ 12.1.044-89 «ССБТ. Пожаровзрывоопасность веществ и материалов. Номенклатура показателей и методы их определения» Режим доступу: WWW. URL: [http://online.budstandart.com/ru/catalog/doc-page?id\\_doc=51048](http://online.budstandart.com/ru/catalog/doc-page?id_doc=51048)

50) НПАОП 0.00-7.15-18 Вимоги щодо безпеки та захисту здоров'я працівників під час роботи з екранними пристроями Міністерство доходів і зборів України Наказ від 05.09.2013 р. № 443 «Про затвердження Примірної інструкції з охорони праці під час експлуатації електронно-обчислювальних машин» Режим доступу: WWW. URL: [http://sop.zp.ua/norm\\_praop\\_0\\_00-7\\_15-18\\_01\\_ua.php](http://sop.zp.ua/norm_praop_0_00-7_15-18_01_ua.php)

51) Нормативно-правові акти з охорони праці. НПАОП 0.00-4.15-98 «Про розробку інструкцій з охорони праці» Режим доступу: WWW. URL: [http://sop.zp.ua/norm\\_praop\\_0\\_00-4\\_15-98\\_01\\_ru.php](http://sop.zp.ua/norm_praop_0_00-4_15-98_01_ru.php)

52) Закон України «Про охорону навколишнього природного середовища» - Режим доступу: <https://zakon.rada.gov.ua/laws/show/1264-12> - 26.06.1991 р.

53) Закони України «Про охорону навколишнього природного середовища» - Режим доступу - <https://zakon.rada.gov.ua/laws/show/4004-12> - 24.02.1994 р.



54) Закон України «Про відходи» - Режим доступу:  
<https://zakon.rada.gov.ua/laws/show/187/98-%D0%B2%D1%80> – 05.03.1998 р.

## ДОДАТОК А.

## Фрагмент лістингу модулю «MyTransform.py»

```

# -*- coding: utf-8 -*-

import cv2
import matplotlib.pyplot as plt
import os
import numpy as np

base_dir = os.getcwd()
data_folder = os.path.join(base_dir, "Dataset")

images_folder = os.path.join(data_folder, "Images")
crops_folder = os.path.join(data_folder, "Crops")
sample_testset = os.path.join(data_folder, "sample_testset")

sample_images = os.path.join(sample_testset, "images")
sample_crops = os.path.join(sample_testset, "crops")

def bruteForceSIFT(img1, img2):

    sift = cv2.xfeatures2d.SIFT_create()

    kp1, des1 = sift.detectAndCompute(img1, None)
    kp2, des2 = sift.detectAndCompute(img2, None)

    bf = cv2.BFMatcher()
    matches = bf.knnMatch(des1, des2, k=2)
    good = []
    print(len(matches))
    for match1, match2 in matches:
        if match1.distance < 0.75 * match2.distance:
            good.append([match1])

    sift_matches = cv2.drawMatchesKnn(img1, kp1, img2, kp2, good, None,
flags=2)

    sift_matches = cv2.cvtColor(sift_matches, cv2.COLOR_BGR2RGB)

    print(len(good))
    plt.imshow(sift_matches)
    plt.show()
    return (kp1, kp2)

def bruteForceORB(img1, img2):
    orb = cv2.ORB_create()
    kp1, des1 = orb.detectAndCompute(img1, None)
    kp2, des2 = orb.detectAndCompute(img2, None)

    bf = cv2.BFMatcher(cv2.NORM_HAMMING, crossCheck=True)

    matches = bf.match(des1, des2)

    matches = sorted(matches, key = lambda x:x.distance)

    orb_matches = cv2.drawMatches(img1, kp1, img2, kp2, matches[:10], None,
flags=cv2.DrawMatchesFlags_NOT_DRAW_SINGLE_POINTS)

    orb_matches = cv2.cvtColor(orb_matches, cv2.COLOR_BGR2RGB)
    plt.imshow(orb_matches)
    plt.show()

```

```

    return matches[:10]

def FLANN(img1, img2):
    FLANN_INDEX_KDTREE = 1
    index_params = dict(algorithm = FLANN_INDEX_KDTREE, trees = 5)
    search_params = dict(checks=10)

    sift = cv2.xfeatures2d.SIFT_create()

    kp1, des1 = sift.detectAndCompute(img1, None)
    kp2, des2 = sift.detectAndCompute(img2, None)

    flann = cv2.FlannBasedMatcher(index_params, search_params)
    matches = flann.knnMatch(des1, des2, k=2)

    matchesMask = [[0,0] for i in range(len(matches))]

    for i, (m,n) in enumerate(matches):
        if m.distance < 0.7*n.distance:
            matchesMask[i]=[1,0]

    draw_params = dict(matchColor = (0,255,0),
                        matchesMask = matchesMask,
                        flags = cv2.DrawMatchesFlags_DEFAULT)

    flann_matches =
cv2.drawMatchesKnn(img1, kp1, img2, kp2, matches, None, **draw_params)

    orb_matches = cv2.cvtColor(flann_matches, cv2.COLOR_BGR2RGB)
    plt.imshow(orb_matches)
    plt.show()

    return matches[:10]

def templateMatching(img1, img2, method):
    w, h, c = img1.shape
    img = img2.copy()

    res = cv2.matchTemplate(img, img1, eval(method))

    min_val, max_val, min_loc, max_loc = cv2.minMaxLoc(res)

    top_left = min_loc

    bottom_right = (top_left[0] + w, top_left[1] + h)

    cv2.rectangle(img, top_left, bottom_right, 255, 2)

    img = cv2.cvtColor(img, cv2.COLOR_BGR2RGB)
    plt.imshow(img)
    plt.show()

    return top_left, bottom_right

def ModifiedFLANN(img1, img2):

    mini_match_count = 10
    FLANN_INDEX_KDTREE = 1
    index_params = dict(algorithm = FLANN_INDEX_KDTREE, trees = 5)
    search_params = dict(checks=10)

    sift = cv2.xfeatures2d.SIFT_create()

    kp1, des1 = sift.detectAndCompute(img1, None)

```

```

kp2, des2 = sift.detectAndCompute(img2, None)

flann = cv2.FlannBasedMatcher(index_params, search_params)
matches = flann.knnMatch(des1, des2, k=2)

good_matches = []
orgBorder = None

for match1, match2 in matches:
    if match1.distance < (0.7 * match2.distance):
        good_matches.append((match1))

if len(good_matches) > mini_match_count:
    cropImg = []
    orgImg = []

    for m in good_matches:
        cropImg.append(kp1[m.queryIdx].pt)
        orgImg.append(kp2[m.trainIdx].pt)

    cropImg, orgImg = np.float32((cropImg, orgImg))

    H, status = cv2.findHomography(cropImg, orgImg, cv2.RANSAC, 3.0)

    h, w, c = img1.shape

    cropBorder = np.float32([[0,0], [0,h-1], [w-1,h-1], [w-1,0]])
    orgBorder = cv2.perspectiveTransform(cropBorder, H)

    cv2.polylines(img2, [np.int32(orgBorder)], True, (0, 255, 0), 5)

else:
    print("No Good Matches Found")

matched = cv2.cvtColor(img2, cv2.COLOR_BGR2RGB)
plt.imshow(matched)
plt.show()

return orgBorder

crop_img = cv2.imread(os.path.join(sample_crops,
                                   "3ef28b3a-cdd5-548b-8134-7c69393c89f2.jpg"))

img = cv2.imread(os.path.join(sample_images,
                               "ea9b55cc-e010-5905-b696-9f8bacc50fe6.jpg"))

plt.imshow(cv2.cvtColor(crop_img, cv2.COLOR_BGR2RGB))
plt.show()
#
=====
#
# matches_sift = bruteForceSIFT(crop_img, img)
#
# matches_orb = bruteForceORB(crop_img, img)
#
# matches_flann = FLANN(crop_img, img)
#
# methods = ['cv2.TM_CCOEFF', 'cv2.TM_CCOEFF_NORMED', 'cv2.TM_CCORR',
#            'cv2.TM_CCORR_NORMED', 'cv2.TM_SQDIFF', 'cv2.TM_SQDIFF_NORMED']
#
# for method in methods:
#     x, y = templateMatching(crop_img, img, method)

```

#

=====

`crop_border = ModifiedFLANN(crop_img, img)`

## ДОДАТОК Б.

## Фрагмент лістингу модулю «MyWalsh.py»

```

import sys
import numpy as np
import struct

def align(x, a):
    return ((x - 1) // a) + 1 * a

class UnsupportedFileFormatError(Exception):
    pass

HEADER_FORMAT = '<2sIHIIIIIIHHIIIIII'

class BMPImage:

    def __init__(self):
        self.__signature = 'BM'
        self.__size = 0
        self.__offset = 54
        self.__headerSize = 40
        self.__width = 0
        self.__height = 0
        self.__planes = 1
        self.__bpp = 24
        self.__compression = 0
        self.__sizeOfData = 0
        self.__horizontalRes = 2835
        self.__verticalRes = 2835
        self.__rawData = None

    def _load_bmp_header(self, file):
        header = file.read(struct.calcsize(HEADER_FORMAT))
        (self.__signature, self.__size, ignored, ignored, self.__offset,
self.__headerSize,
        self.__width, self.__height, self.__planes, self.__bpp,
self.__compression,
        self.__sizeOfData, self.__horizontalRes, self.__verticalRes,
ignored, ignored) = \
            struct.unpack(HEADER_FORMAT, header)
        if (self.__signature, self.__planes, self.__bpp, self.__compression)
!= ('BM', 1, 24, 0):
            raise UnsupportedFileFormatError
#         print 'Size of data: ' + str(self.__sizeOfData)

    def _load_bmp_data(self, file):
        file.seek(self.__offset)
        self.__rawData = []
        for j in xrange(0, self.__height):
            line = file.read(align(self.__width * 3, 4))
            for i in xrange(0, self.__width * 3, 3):
                self.__rawData.append(struct.unpack_from('<BBB', line, i))

    def load(self, filename):
        f = sys.stdin
        if filename and isinstance(filename, str):
            f = open(filename, 'rb')
        with f as file:
            self._load_bmp_header(file)
            self._load_bmp_data(file)

```

```

def _write_header(self, file):
    ignored = 0
    header = (
        self.__signature, self.__size, ignored, ignored, self.__offset,
self.__headerSize,
        self.__width, self.__height, self.__planes, self.__bpp,
self.__compression,
        self.__sizeOfData, self.__horizontalRes, self.__verticalRes,
ignored, ignored
    )
    raw_header = struct.pack(HEADER_FORMAT, *header)
    file.write(raw_header)

def _write_bmp_data(self, file):
    padding = align(self.__width * 3, 4) - self.__width * 3
    print 'Padding: ' + str(padding)
    counter = 0
    for v in self.__rawData:
        r, g, b = v
        file.write(struct.pack('<BBB', r, g, b))
        counter += 1
        if counter == self.__width:
            counter = 0
            for i in xrange(0, padding):
                file.write(struct.pack('<B', 0))

def save(self, filename):
    with open(filename, 'wb') as file:
        self._write_header(file)
        self._write_bmp_data(file)

def get_dimensions(self):
    return self.__width, self.__height

def set_dimensions(self, width, height):
    self.__width = width
    self.__height = height
    self.__sizeOfData = width * height * 3

def get_raw_data(self):
    return self.__rawData

def set_raw_data(self, new_data):
    self.__rawData = new_data

class CustomizableImage:

    def __init__(self):
        self.__width = 0
        self.__height = 0

        # (originalBlockSize, packedBlockSize, numberOfBlocks)
        self.__yDescription = None
        self.__cbDescription = None
        self.__crDescription = None

        self.__yData = None
        self.__cbData = None
        self.__crData = None

    def _read_header(self, file):
        self.__width, self.__height = struct.unpack('<II',
file.read(struct.calcsize('<II')))
        self.__yDescription = struct.unpack('<HHH',
file.read(struct.calcsize('<HHH')))

```

```

        self.__cbDescription = struct.unpack('<HHH',
file.read(struct.calcsize('<HHH')))
        self.__crDescription = struct.unpack('<HHH',
file.read(struct.calcsize('<HHH')))

    @staticmethod
    def _read_blocks(file, description):
        blocks = []
        original_block_size = description[0]
        packed_block_size = description[1]
        number_of_blocks = description[2]
        block_length = packed_block_size * packed_block_size
        pattern = '<' + 'h' * block_length
        for i in xrange(0, number_of_blocks):
            data = struct.unpack(pattern,
file.read(struct.calcsize(pattern)))
            block = np.matrix(np.zeros((original_block_size,
original_block_size)))
            block[0:packed_block_size, 0:packed_block_size] =
np.matrix(data).reshape(packed_block_size,
packed_block_size)
            blocks.append(block)
        return blocks

    @staticmethod
    def load(filename):
        f = sys.stdin
        if filename and isinstance(filename, str):
            f = open(filename, 'rb')
        with f as file:
            image = CustomizableImage()
            image._read_header(file)
            if image.__yDescription[2] > 0:
                image.__yData = image._read_blocks(file,
image.__yDescription)
            if image.__cbDescription[2] > 0:
                image.__cbData = image._read_blocks(file,
image.__cbDescription)
            if image.__crDescription[2] > 0:
                image.__crData = image._read_blocks(file,
image.__crDescription)
            return image
        def get_y_data(self):
            return self.__yData
        def get_cb_data(self):
            return self.__cbData
        def get_cr_data(self):
            return self.__crData
        def get_dimensions(self):
            return self.__width, self.__height
        def set_dimensions(self, width, height):
            self.__width = width
            self.__height = height
        def set_descriptions(self, y_description, cb_description,
cr_description):
            self.__yDescription = y_description
            self.__cbDescription = cb_description
            self.__crDescription = cr_description

        def set_data(self, y_data, cb_data, cr_data):
            y_packed_block_size = self.__yDescription[1]
            self.__yData = map(lambda x: x[0:y_packed_block_size,
0:y_packed_block_size], y_data)
            cb_packed_block_size = self.__cbDescription[1]
            self.__cbData = map(lambda x: x[0:cb_packed_block_size,
0:cb_packed_block_size], cb_data)

```



```

    cr_packed_block_size = self.__crDescription[1]
    self.__crData = map(lambda x: x[0:cr_packed_block_size,
0:cr_packed_block_size], cr_data)

def _write_header(self, file):
    file.write(struct.pack('<II', self.__width, self.__height))
    file.write(struct.pack('<HHH', *self.__yDescription))
    file.write(struct.pack('<HHH', *self.__cbDescription))
    file.write(struct.pack('<HHH', *self.__crDescription))

@staticmethod
def _write_blocks(file, blocks):
    for block in blocks:
        data = np.array(block).reshape(-1).tolist()
        file.write(struct.pack('<' + 'h' * len(data), *data))

def save(self, filename):
    with open(filename, 'wb') as file:
        self._write_header(file)
        if self.__yData:
            self._write_blocks(file, self.__yData)
        if self.__cbData:
            self._write_blocks(file, self.__cbData)
        if self.__crData:
            self._write_blocks(file, self.__crData)

```

## ДОДАТОК В.

### Електронні плакати

*Міністерство освіти і науки України  
Східноукраїнський національний університет  
ім. В. Даля  
Кафедра комп'ютерних наук та інженерії*

*Магістерська робота  
«Моделі і методи класифікації  
зображень для систем  
комп'ютерного зору»*

*Виконав: ст.гр. КН-19дм*

*Коробка Я.О.*

*Керівник: Недзельський Д.О.*

## МЕТА ДИПЛОМНОЇ РОБОТИ

**Об'єктом дослідження** є методи розпізнавання зображень у системах комп'ютерного зору.

**Метою дослідження** є розроблення методів зіставлення структурних описів зображень на основі апарату мультимножин та застосування результативних трансформацій описів.

**Основним завданням** прийнято розробку методів обчислення релевантності описів зображень із використанням апарату мультимножин, проведення експериментальних досліджень розроблених методів, порівняння результатів з відомими методами для конкретних баз візуальних даних. Для отримання бінарних дескрипторів ключових точок застосовано детектор ORB.

## АНАЛІЗ МЕТОДІВ ІНВАРІАНТНОГО РОЗПІЗНАВАННЯ ЗОБРАЖЕНЬ



## АНАЛІЗ МЕТОДІВ ІНВАРІАНТНОГО РОЗПІЗНАВАННЯ ЗОБРАЖЕНЬ

$$\mathfrak{R}: Z \rightarrow \{Z(j)\} \quad (1)$$

$$\alpha(Z) = \arg \operatorname{opt}_{j \in \{1, \dots, J\}} \mathfrak{S}[Z, Z(j)] \quad (2)$$

$\mathfrak{S} \in \mathbb{R}^1$  – міра подібності;

$\mathfrak{C}(Z)$  – функція отримання класу об'єкту.

$$\rho_{\mathfrak{H}}(A, B) = \max[\rho(A, B), \rho(B, A)] \quad (3)$$

## ПОСТАНОВКА ЗАДАЧІ ДОСЛІДЖЕННЯ

Для реалізації мети слід **вирішити** такі задачі:

- зниження надмірності множини характерних ознак (МХО) шляхом створення неповного або стисненого опису об'єкта, і як результат, скорочення обсягу структурного опису без втрати якості розпізнавання;
- побудова нових покращених мір подібності на основі найбільш значущих елементів опису, що підвищує швидкодію розпізнавання без зниження перешкодозахищеності;
- підвищення ефективності розпізнавання шляхом застосування для зіставлення МХО апарату мультимножин, що інтегрує інформацію про описи.

## СТИСНЕННЯ ОПИСУ НА ОСНОВІ ЗАСТОСУВАННЯ ОРТОГОНАЛЬНОГО БАЗИСУ ФУНКЦІЙ УОЛША

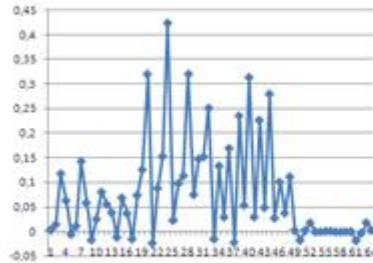
$$A = \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \quad (1)$$

$$U = Z * A \quad (2)$$

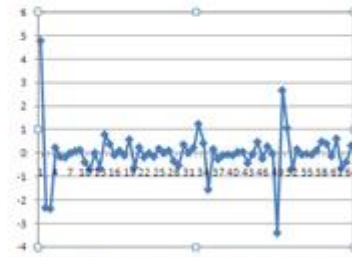
де A – матриця Адамара, що включає вектори ФУ  $w_1, \dots, w_2$

$$U = \begin{bmatrix} u_1^1 \dots u_n^1 \\ u_1^2 \dots u_n^2 \\ \dots \\ u_1^s \dots u_n^s \end{bmatrix} = ZA = \begin{bmatrix} z_1^1 \dots z_n^1 \\ z_1^2 \dots z_n^2 \\ \dots \\ z_1^s \dots z_n^s \end{bmatrix} \times \frac{1}{n} \begin{bmatrix} a_{11} \dots a_{1n} \\ a_{21} \dots a_{2n} \\ \dots \\ a_{n1} \dots a_{nm} \end{bmatrix} \quad (3)$$

## ДОСЛІДЖЕННЯ ВЛАСТИВОСТЕЙ ТРАНСФОРМОВАНИХ ОПИСІВ



Дескриптор ХО, сформований методом SURF



Дескриптор ХО після перетворення Уолша

## ОПТИМІЗАЦІЯ ПАРАМЕТРА ПОРОГА ПРИ ЗІСТАВЛЕННІ ОПИСІВ ОБ'ЄКТІВ

$$\max A_1(\delta_z) = \frac{1}{c} \sum_{i=1}^c h_{ii} / \max(h_i^*)$$

$$\min A_2(\delta_z) = \frac{1}{c} \sum_{i=1}^c 1(\max_{j=1, \dots, c-1} h_{ij} \geq 0,5 h_{ii})$$

де  $A_1$  – середнє по числу еталонів ставлення діагональних елементів  $H$  до їх максимально можливих значення  $\max(h_i^*)$ ,

$A_2$  – частка числа рядків матриці  $H$ , в яких найближчий до діагонального значенням максимум більше половини діагонального елемента  $h_{ii}$ ,

$c$  – число еталонів бази.

$$1(\max_{j=1, \dots, c-1} h_{ij} \geq 0,5 h_{ii}) = \begin{cases} 1, & \max_{j=1, \dots, c-1} h_{ij} \geq 0,5 h_{ii}, \\ 0, & \max_{j=1, \dots, c-1} h_{ij} < 0,5 h_{ii} \end{cases}$$



## РЕЗУЛЬТАТИ ЕКСПЕРИМЕНТАЛЬНИХ ДОСЛІДЖЕНЬ



## ВИСНОВКИ

1. Встановлено, що аналіз застосовуваних ОО і виявлення особливостей бази зображень поряд з застосуванням методом значно впливають на результат розпізнавання. Застосування процедур перетворення опису або переходу в новий простір ознак сприяє поліпшенню характеристик розпізнавання;
2. Обраний метод побудови описів на основі функцій Уолша дозволяє проводити розпізнавання об'єктів з високою швидкістю без зниження перешкодозахищеності;
3. Розроблена модель стиснення ОО за результатами експерименту дозволяє зберегти безпомилкове розпізнавання на рівні 40% від усього числа точок опису. Це допускає рівень стиснення вихідного простору ознак опису в 2,5 рази;
4. Удосконалена модель побудови мір подібності на основі принципу найближчих сусідів забезпечує високу ймовірність правильної класифікації. Модель подібності описів забезпечує порівнянню з традиційними підходами перешкодозахищеність;
5. Реалізація запропонованих моделей і методів дозволяє вирішити ряд прикладних задач. Досягнуто необхідний рівень достовірності та швидкості при класифікації зображень риб.