

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет Інформаційно-аналітичних технологій та менеджменту
(повна назва)
Кафедра Інформатики
(повна назва)

АТЕСТАЦІЙНА РОБОТА
Пояснювальна записка

рівень вищої освіти другий (магістерський)
ДОСЛІДЖЕННЯ ГІБРИДНИХ МЕТОДІВ ДЛЯ КЛАСИФІКАЦІЇ
СКЛАДНОСТРУКТУРОВАНИХ ЗОБРАЖЕНЬ
(тема)

Виконав:
студент 2 курсу, групи ІНФМ-18-2
Зеленський М.О.
(прізвище, ініціали)
Спеціальності 122 Комп'ютерні науки
(код і повна назва спеціальності)
Тип програми освітньо-професійна
(освітньо-професійна або освітньо-наукова)
Освітня програма Інформатика
(повна назва освітньої програми)
Керівник доц. Творошенко І.С.
(посада, прізвище, ініціали)

Допускається до захисту

Зав. кафедри _____
(підпис) Путятін Є.П.
(прізвище, ініціали)

2019 р.

Харківський національний університет радіоелектроніки

Факультет Інформаційно-аналітичних технологій та менеджменту
(повна назва)Кафедра Інформатики
(повна назва)Рівень вищої освіти другий (магістерський)Спеціальність 122 Комп'ютерні науки
(код і повна назва)Тип програми освітньо-професійна
(освітньо-професійна або освітньо-наукова)Освітня програма Інформатика
(повна назва освітньої програми)

ЗАТВЕРДЖУЮ:

Зав. кафедри _____
(підпис)

« ____ » _____ 20 ____ р.

ЗАВДАННЯ
НА АТЕСТАЦІЙНУ РОБОТУстудентові Зеленському Михайлу Олександровичу
(прізвище, ім'я, по батькові)1. Тема роботи Дослідження гібридних методів для класифікації складноструктурованих зображень

затверджена наказом по університету від 21 жовтня 2019 року № 1506Ст

2. Термін подання студентом роботи до екзаменаційної комісії 28 листопада 2019 р.3. Вихідні дані до роботи науково-методична література, дані інтернет-мережі, методи та засоби класифікації складноструктурованих зображень: інтелектуальні оператори сегментації та методологія спектрального опису кривих складної форми складноструктурованих зображень; мережеві інтелектуальні технології для класифікації сегментів на складноструктурованих зображеннях; гібридні методи для класифікації складноструктурованих зображень.4. Перелік питань, що потрібно опрацювати в роботі аналіз методів та засобів класифікації складноструктурованих зображень, інтелектуальні оператори сегментації та методологія спектрального опису кривих складної форми складноструктурованих зображень, мережеві інтелектуальні технології для класифікації сегментів на складноструктурованих зображеннях, дослідження гібридних методів для класифікації складноструктурованих зображень.

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій (слайдів) Актуальність дослідження, мета та об'єкт дослідження, постановка задачі дослідження, аналіз предметної області, методи вирішення поставлених задач, вихідні данні, дослідження запропонованих методів при вирішенні поставлених задач, тестування програмного забезпечення, перспективи подальших досліджень, висновки, апробація результатів дослідження.

6. Консультанти розділів роботи (п.6 включається до завдання за наявності консультантів згідно з наказом, зазначеним у п.1)

Найменування розділу	Консультант (посада, прізвище, ім'я, по батькові)	Позначка консультанта про виконання розділу	
		підпис	дата

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів роботи	Терміни виконання етапів роботи	Примітка
1	Отримання завдання на атестаційну роботу	21.10.2019	
2	Аналіз завдання, підбір літератури	22.10.19–23.10.19	
3	Аналіз літератури з досліджуваної проблеми	24.10.19–26.10.19	
4	Аналіз технічних засобів	27.10.19–29.10.19	
5	Розробка методів	30.10.19–14.11.19	
6	Програмна реалізація	15.11.19–17.11.19	
7	Оформлення пояснювальної записки	18.11.19–19.11.19	
8	Перевірка на плагіат	20.11.19	
9	Рецензування	21.11.19	
10	Підготовка презентації та доповіді	22.11.19	
11	Занесення роботи в електронний архів	23.11.19	
12	Попередній захист атестаційної роботи	25.11.19	

Дата видачі завдання 21 жовтня 2019 року

Студент _____
(підпис)

Керівник роботи _____
(підпис)

доц. Творошенко І.С.
(посада, прізвище, ініціали)

РЕФЕРАТ/ABSTRACT

Пояснювальна записка до атестаційної роботи: 110 с., 5 табл., 33 рис., 1 дод., 40 джерел.

СКЛАДНОСТРУКТУРОВАНІ ЗОБРАЖЕННЯ, КЛАСИФІКАЦІЯ, ДЕСКРИПТОРИ ФУР'Є, АНАЛІЗ КОРДОНІВ, МЕТОД СЕГМЕНТАЦІЇ, НЕЙРОМЕРЕЖЕВІ СТРУКТУРИ.

Метою дослідження є виявлення ефективних гібридних методів для класифікації складноструктурованих зображень.

Об'єктом дослідження є складноструктуровані зображення.

Використано методи сегментації, що засновані на кластеризації на градієнтних операторах, модельна сегментація, узагальнений метод сегментації. Досліджено та протестовано методи для класифікації складноструктурованих зображень.

У результаті роботи здійснена програмна реалізація різних методів, як окремих компонент.

Результати атестаційної роботи апробовано у вигляді тез доповіді на міжнародній конференції.

COMPLEX STRUCTURED IMAGES, CLASSIFICATION, FURIE DESCRIPTORS, BORDER ANALYSIS, SEGMENTATION ALGORITHM, NEURIET STRUCTURES.

The purpose of the study is to identify effective hybrid methods for the classification of complex structured images.

The object of the study is complex structured images.

Used segmentation methods based on clustering by gradient operators, model segmentation, generalized segmentation algorithm. Methods for classifying complex structured images are investigated and tested.

Because of the work, the software implementation of various methods as separate components is carried out.

The results of the performance appraisal were tested in the form of abstracts at an international conference.

ЗМІСТ

Перелік умовних позначень, символів, одиниць, скорочень і термінів	7
Вступ.....	8
1 Аналіз методів та засобів класифікації складноструктурованих зображень	9
1.1 Опис складних об'єктів та методів їх подання	9
1.2 Аналіз існуючих математичних моделей зображень	15
1.3 Аналіз методів сегментації складноструктурованих зображень ...	20
1.3.1 Порогова сегментація	24
1.3.2 Методи сегментації, що засновані на кластеризації.....	26
1.3.3 Методи сегментації, що засновані на градієнтних операторах	27
1.3.4 Модельна сегментація.	31
1.4 Опис процедур та методів класифікації зображень	36
1.5 Постановка задачі дослідження.....	39
2 Інтелектуальні оператори сегментації та методологія спектрального опису кривих складної форми складноструктурованих зображень.....	41
2.1 Узагальнений метод сегментації складноструктурованих зображень	41
2.2 Синтез морфологічного оператора для препарування зображень.	42
2.3 Особливості обробки бінарних зображень.....	52
2.3.1 Нечітка модель формування бінарного зображення границі сегменту	52
2.3.2 Метод морфологічної обробки бінарних зображень.....	65
2.4 Дослідження методів спектрального уявлення контурів мінімальної товщини.....	71
2.5 Вибір та оптимізація кількості дескрипторів Фур'є при аналізі кордонів сегментів складноструктурованих зображень	76

3 Мережеві інтелектуальні технології для класифікації сегментів на складноструктурованих зображеннях.....	82
3.1 Універсальна мережева структура для задачі класифікації	82
3.2 Нейромережеві структури для сегментації зображень за параметрами пікселів та їх околиць	90
4 Дослідження гібридних методів для класифікації складноструктурованих зображень	96
4.1 Вибір інструментального засобу для вирішення задач.....	96
4.2 Вихідні дані для дослідження.....	97
4.3 Тестування гібридних методів для класифікації складноструктурованих зображень	98
4.4 Результати оцінювання ефективності гібридних методів для класифікації складноструктурованих зображень	101
4.5 Перспективи подальших досліджень.....	102
Висновки	103
Перелік джерел посилання	104
Додаток А Лістинги програми за різними методами	108

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ

ОПР – особа, яка приймає рішення.

ЕМ-метод – Expectation-maximization (EM) method

СТЗ – системи технічного зору

RGB – Red-Green-Blue – основний спектр кольорів

ML – Machine Learning

AI – Artificial Intelligence

НБС – неблокова структура

БС – блокова структура

ВСТУП

Створення і використання систем штучного інтелекту вимагає спеціального теоретичного осмислення загальних закономірностей побудови та функціонування об'єктів різної природи.

Сучасна наука шукає способи і методи визначення в складних системах властивостей окремих об'єктів, їх суворої формалізації на основі останніх досягнень математики, логіки, кібернетики, інформатики та інших точних наук. Аналіз складних об'єктів ґрунтується на методології системного підходу. Системний підхід, як правило, спирається на математичне моделювання з використанням теорії подібності, теорії наукового експерименту, математичної статистики, теорії методів і ряду інших фундаментальних класичних теорій.

Більшість задач класифікації зображень зводиться до вибору однієї з двох альтернатив: «так» – «ні» або «зображення / сегмент відноситься до заданого класу» – «зображення / сегмент не відноситься до заданого класу». Ця добре відома задача класифікації двухальтернативної вибірки.

Саме через це у даній роботі необхідно проілюструвати методи та основні принципи обробки складноструктурованих зображень, визначити та провести аналіз математичних моделей.

Протягом дослідження необхідно визначити закономірності серед методів обробки зображень, а також визначити область застосувань дескрипторів Фур'є у сфері обробки даного виду зображень.

1 АНАЛІЗ МЕТОДІВ ТА ЗАСОБІВ КЛАСИФІКАЦІЇ СКЛАДНОСТРУКТУРОВАНИХ ЗОБРАЖЕНЬ

1.1 Опис складних об'єктів та методів їх подання

У практиці штучного інтелекту доводиться мати справу з такими поняттями як об'єкт і система. Основними методологічними категоріями опису складних об'єктів є поняття: об'єкт, клас, ставлення (зв'язок), система. У найбільш загальному уявленні об'єкт – це все, що особа, яка приймає рішення (ОПР), розрізняє як щось ціле і реально існуюче, або виникає в його свідомості і володіє властивостями, значення яких дозволяють ОПР однозначно розпізнавати це щось.

Під системою будемо розуміти множина елементів, що знаходяться у відносинах і зв'язках один з одним, яке утворює певну цілісність, єдність. Відмінність системи від об'єкта полягає в тому, що на систему накладаються додаткові обмеження, відповідні її 19 функціональному стану, тобто здатності виконувати певний вид діяльності.

Таким чином, об'єкт сприймається як ціле незалежно від його функціонального стану або від зв'язків між елементами, тоді як система сприймається як система тільки при наявності всіх зв'язків, які забезпечують її функціональний стан. Складна система – система, що складається з множини взаємодіючих складових (підсистем), внаслідок чого складна система набуває нових властивостей, які відсутні на підсистемному рівні і не можуть бути зведені до властивостей підсистемного рівня.

У теорії штучного інтелекту поняття «об'єкт» грає роль фундаментального вихідного поняття, через яке визначаються всі основні поняття. Воно не може бути виведено з будь-яких більш загальних понять шляхом успадкування їх властивостей. Будь-які об'єкти визначаються шляхом деталізації і конкретизації атрибутів даного поняття:

- наявність унікального імені;

- визначеність призначення;
- наявність внутрішньої структури;
- наявність особливих властивостей або ознак, що дозволяють;
- ідентифікувати (дізнаватися) даний об'єкт;
- знаходження в певному просторі;
- знаходження в певному стані;
- можливість зміни положення і стану об'єкта;
- наявність відносин і зв'язків з іншими об'єктами.

Допускається також доповнювати, конкретизувати або змінювати цей перелік атрибутів в залежності від природи визначених об'єктів. Взагалі кажучи, атрибут є суттєва деталь, параметр, властивість або функція визначається об'єкта. Іншими словами, атрибут є об'єкт, службовець для визначення іншого об'єкта, і він так само, як і будь-який інший об'єкт, має свої атрибути.

Залежно від призначення атрибути об'єктів можна розділити на:

- ідентифікатори;
- параметри;
- методи взаємодії і функції, які визначають динаміку внутрішніх;
- станів об'єктів і зовнішній ефект, що виникає в результаті цієї динаміки.

Межі між перерахованими групами атрибутів досить умовні. Призначення атрибутів при описі об'єкта є прерогативою дослідника і повністю визначається цілями і завданнями дослідження. Об'єкти, визначені на одних і тих же атрибутах з різними значеннями, відносять до одного класу об'єктів. Процес об'єднання об'єктів в класи відповідно до їх властивостями називається класифікацією. Насправді мислення людини засноване на абстракціях (поняттях), що позначають класи різних об'єктів. При аналізі складних систем дослідник (розробник) повинен вибирати необхідні йому екземпляри об'єктів із заданого класу об'єктів шляхом конкретизації значень їх атрибутів (ідентифікації).

Складність вибору необхідного варіанту (примірника об'єктів) оцінюється величиною різноманітності (потужністю) даного класу об'єктів. «Клас» – термін, що вживається в математиці в основному як синонім терміну «множина» для позначення довільних сукупностей об'єктів, що володіють якою-небудь певним властивістю, або ознакою (наприклад, в алгебрі – класи еквівалентності щодо даного відносини еквівалентності).

Іноді класами воліють називати сукупності, елементами яких є множини. У деяких випадках під впливом аксіоматичної теорії множин термін «клас» застосовується для того, щоб підкреслити, що дана сукупність виявляється власне класом, а не множиною у вузькому сенсі (наприклад, в алгебрі – примітивні класи універсальних алгебр, звані також різноманіттям). Відомо, що ряд базисних визначень, пов'язаних з теорією множин, носять інтуїтивний характер і в зв'язку з цим виникає ряд протиріч, на які одні дослідники не звертають уваги, інші – загострюють ці проблеми.

У теорії множин розвиваються два підходи, які відомі в літературі як канторовський і расселовський.

Основні суперечності в цих підходах пов'язані з трактуванням цілого. При розподілі цілого – об'єкта на частини різними способами отримуємо різні непересічні множини або це має бути одне множина? Поки однозначну відповідь на це питання не сформульований, точніше різними вченими він формулюється по-різному.

У прикладній області об'єкт – це множина атрибутів – ознак або предикторів – провісників, які його визначають і в тій чи іншій мірі характеризують клас, що відповідає расселовського поняттю класу, як набору елементів. Тобто предиктори, а точніше їх значення, характеризують ступінь належності об'єкта до класу. Ступінь приналежності описується характеристичною функцією. Пізніше, коли була побудована теорія нечітких моделей Заде [1], характеристическую функцію в нечіткому аналізі стали називати функцією приналежності, мабуть тому, що у характеристичної функції багато трактувань в різних областях знань.

Теоретико-множинні операції над класами визначаються так само, як і над множинами. У загальному випадку класифікація об'єктів може проводитися за кількома визначальними ознаками (атрибутами об'єктів). Наприклад, людей можна класифікувати за іменем, статтю, днем народження, освіти, професії і т.д. Кожна конкретна людина даного класу буде визначатися (ідентифікуватися) відповідним унікальним, тільки йому властивим, набором значень даних атрибутів, який в подальшому будемо називати кодом або ідентифікатором даного об'єкта.

Нехай $A = \{A_1, \dots, A_N\}$ – множина незалежних атрибутів, що характеризують певну множину об'єктів $O = \{O_1, O_2, \dots, O_m\}$. Поставимо в однозначна відповідність кожному атрибуту A_i багатозначний предиктор P_i , що визначає множина значень $E_i = \{a_1, \dots, a_k\}$ і даного атрибута A_i так, що

$$\forall t \mid P_i = (A_i(t) = a_e \in E_i), \quad l = \overline{1, k}. \quad (1.1)$$

Іншими словами, в будь-який момент часу t предиктор P_i може приймати тільки одне якесь значення a_e з множини E_i . Тоді клас K є абстракція, яка визначається як декартовий добуток $\{E_1 \times \dots \times E_N\} = \{K_1, K_2, \dots, K_m\}$, K_j – код або ідентифікатор конкретного об'єкта O_j . За цим визначенням символи O_j і K_j позначають поняття, що перебувають у відношенні взаємно однозначної відповідності, $O_j \Leftrightarrow K_j$. Код K_j – є символічна (абстрактна) модель об'єкта O_j . Конкретний об'єкт O_j – це екземпляр класу об'єктів з множини об'єктів $O = \{O_1, O_2, \dots, O_m\}$, $O_j \in O$, що складає даний клас.

У процесі класифікації класу об'єктів O за певними правилами ставиться у відповідність клас представляють об'єктів (моделей) – K .

Дане визначення класу, по суті, еквівалентно теоретико-множинному поняттю універсуму U . Це означає, що клас як математичний формалізм має всі властивості формалізму універсум. Тому для опису і аналізу відносин між об'єктами і підмножинами об'єктів одного класу можна успішно застосовувати обчислення узагальнених кодів підмножин.

При цьому атрибути $A_1, \dots, A_n$ об'єктів класу O можна розглядати як осі класифікаційного простору (простору ідентифікаторів) об'єктів $\{O_1, O_2, \dots, O_m\}$, де кожен конкретний об'єкт O_j представлений крапкою з координатами K_j , як показано на рисунку 1.1.

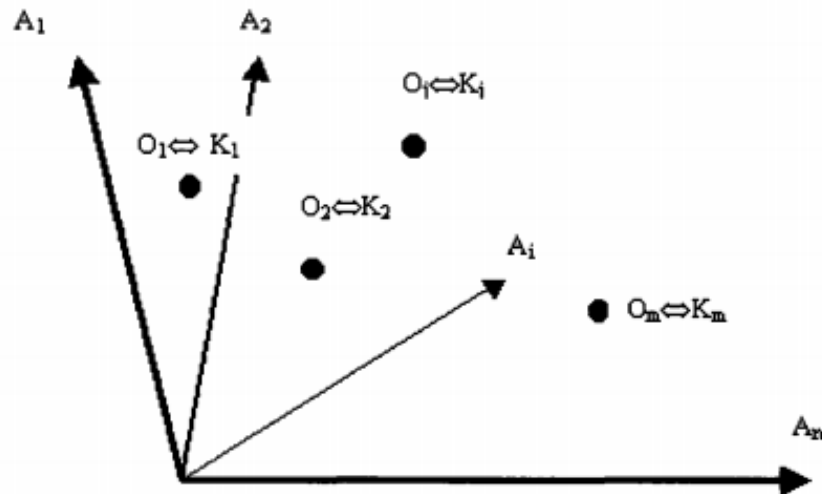


Рисунок 1.1 – Осі класифікаційного простору об'єктів

Класифікаційні атрибути об'єктів служать для ідентифікації конкретних об'єктів. Поряд з класифікаційними атрибутами, як правило, не міняють своїх значень з плином часу, об'єкти мають також атрибути, які можуть змінювати свої значення з часом під впливом зовнішніх і внутрішніх факторів, тобто, вони є функціями часу та інших факторів $A_i^* = f_i(t; X)$. Символом X тут позначена множина всіх можливих факторів, що впливають на об'єкт класу O .

Набір значень всіх атрибутів об'єкта в певний момент часу t називається станом даного об'єкта. Атрибути $A_i^* = f_i(t, X)$ характеризують стан об'єкта класу O , тому вони називаються параметрами даного об'єкта. Кожен об'єкт O_j , що належить до певного класу об'єктів, має свій ідентифікатор K_j в універсумі

K і володіє своєю власною поведінкою, яке характеризується функцією $q_j = f_j(t, K_j, X), j = 1, \dots, m$.

Поведінка об'єкта O_j залежить як від власних параметрів, так і від зовнішніх факторів, в тому числі і від впливів з боку інших об'єктів, які можуть належати різним класам. Згідно з введеними позначеннями, будь-який об'єкт O_j класу O може бути представлений набором значень своїх координат.

При цьому поведінка об'єкта O_j є процес

$$\tilde{O}_j = F_j(K_j, Q_j, X, t), \quad (1.2)$$

де X – базис факторного простору. Фактор часу t винесено з множини X в силу його особливої значущості в процесі поведінки об'єкта.

У рівнянні (1.2) множини всіх атрибутів об'єкта O_j розбивається на дві підмножини: K і Q . Підмножина K містить статичні атрибути об'єкта O , а підмножина Q містить атрибути, що характеризують поведінку даного об'єкта, тобто $A = K \cup Q$.

Значення атрибутів $A_i^* = f_i(t, X)$ також можуть бути сформульовані за допомогою предикторів і об'єднані в послідовність змінних, яка є базою координат простору станів певного об'єкта.

У складних об'єктах кожен предиктор (атрибут) сам є об'єктом. Найбільш типовий приклад такого атрибута – тиражування проекції об'єкта на будь-яку з координат, яка, в окремому випадку, спеціально вводиться для цих цілей. Якщо вводиться тимчасовий атрибут, то проектуючи об'єкт на тимчасову координату, отримуємо об'єкти, які є тимчасовими атрибутами вихідного об'єкта. У реальних системах час дискретно, то j -й об'єкт класу O описується кортежем множин

$$\tilde{O}_j = \{F_j(K_j, Q_j(t_s), X, t)\}, \quad (1.3)$$

де t_s – кортеж з S тимчасових дискретів спостереження об'єкта $\tilde{O}_{j,s} = \overline{I, S}$.

Рівняння (1.3) описує матрицю, число елементів в рядку якої відповідає числу елементів у множині A , а число рядків визначається числом дискретів S . Якщо атрибути – об'єкти в (1.3) описуються об'єктами, що мають не більше трьох атрибутів, то ця матриця може бути представлена у вигляді растрового зображення.

Якщо число атрибутів в об'єктах більше трьох, то тоді об'єкт необхідно спроектувати на певну вісь або вибрати функціональне перетворення, що приводить до скорочення атрибутів, що описують об'єкт.

Таким чином, для опису складних об'єктів, як атрибути використовуються растрові зображення, одержувані шляхом проєкції об'єкта на тимчасову вісь або на осі, що визначають інші атрибути.

1.2 Аналіз існуючих математичних моделей зображень

З математичної точки зору під зображенням розуміють числову функцію f , задану на обмеженій підмножині X , площині R^2 . При цьому область X називають полем зору, а значення $f(x)$ функції f в точці $x \in X$ – яскравістю в точці x поля зору X [2].

На практиці для представлення зображень в якості поля зору X вибирають набір вузлів прямокутної сітки, а яскравість в кожному вузлі сітки визначається цілим числом, що належить $[0, 255]$. Однак при додаванні або відніманні двох зображень результат може вийти за межі $[0, 255]$. Тому прийнято вважати, що поле зору X – довільно обмежена підмножина площині R^2 , а яскравість зображення в кожній точці поля зору X приймає числове значення: $f(x) \in (-\infty, \infty)$, $x \in X$. Лінійні операції додавання зображень і множення зображень на число мають вигляд:

$$(f + g)(x) = f(x) + g(x),$$

$$(a \cdot f)(x) = a \cdot f(x), \quad x \in X. \quad (1.4)$$

При виконанні умови (1.4) множина всіх зображень, заданих на поле зору X , є лінійним простором.

Як вимірювання (оцінки) зображень можуть використовуватися деякі лінійні комбінації яркостей в множині точок (вузлів), які належать полю зору X .

Таким чином, зображення може бути представлене як вимірна множина. Визначивши міру на цій множині, можемо ввести множина зображень. Якщо ця множина задовольняє аксіоматиці простору, то множина зображень може розглядатися як простір зображень.

Залежно від того, як визначено відстань між зображеннями, розглядаються різні метричні простори зображень. При аналізі та класифікації зображень, що складно структуруються найбільш часто використовують лінійний евклідов простір $L^2(X)$ зі скалярним добутком

$$(f, g) = \int_X f(x) \cdot g(x) d\mu(x). \quad (1.5)$$

Якщо на множині X задано сітка вузлів і яка вважає міра, то для обчислення норми, скалярного твору і відстані між зображеннями достатньо знати значення зображень в вузлах сітки. У цьому випадку будь-яке зображення $f \in L^2_\mu(X)$ будемо розглядати як вектор n -мірного простору з координатами

$$(f(x_1), \dots, f(x_n)) \in \mathbb{R}^n. \quad (1.6)$$

Скалярний добуток будь-яких двох зображень f і g визначається у вигляді:

$$(f, g) = \sum_{i=1}^n f(x_i) \cdot g(x_i). \quad (1.7)$$

При цьому відстань між зображеннями f і g визначається нормою різниці $f - g$:

$$\|f - g\| = (\sum_{i=1}^n (f(x_i) - g(x_i))^p)^{1/p}. \quad (1.8)$$

Норма зображення f визначається у вигляді:

$$\|f\| = (\sum_{i=1}^n f^2(x_i))^{1/2}. \quad (1.9)$$

При визначенні відстані між двома зображеннями згідно (1.8) вважають, що вклад яркостної характеристики пікселя в міру не залежить від координати (індексу) пікселя i .

При дослідженні складноструктурованих зображень це положення не виконується, тому в формулу (1.8) необхідно ввести відповідні вагові коефіцієнти. При цьому, в залежності від структури (класу) зображень, коефіцієнти можуть бути однаковими для простору $L^2_\mu(X)$, тобто не залежати від f та g , а можуть змінюватися від одного зображення до іншого.

Основна ідея аналізу складноструктурованих зображень полягає в математичному дослідженні моделі формування зображень з метою пошуку структури зображення в інваріантної до тих чи інших перетворень зареєстрованого зображення і параметрам моделі. Форма зображення визначається як максимальне число інваріантів таких перетворень або іншими словами, форма – множина зображень фіксованої сцени при всіх можливих умовах реєстрації [2]. Форма зображення є центральним поняттям морфологічного аналізу і пов'язана з моделлю реєстрації зображення об'єкта або сцени. Якщо розглядається півтонове зображення, отримане від об'єкта, поверхня якого складається з ділянок оптично однорідних площин, освітлюваних однорідним світловим потоком, то кожна ділянка такій площині на зображенні об'єкта буде мати однорідну яскравість.

Зображення такої ділянки можна розглядати як функцію $f(\cdot)$ задану на поле зору $X \in R^2$, приймаючи числові значення і перед ставиме у вигляді:

$$f(x) = \sum_{i=1}^N c_i \chi_i(x), \quad x \in X. \quad (1.10)$$

де $\chi_i(x)$ – індикаторна функція множини A_i (поля зору X , всі точки якого мають однакову яскравість c_i , $i = 1, \dots, N$ і визначується у вигляді:

$$\chi_i(x) = \begin{cases} 1, & x \in A_i \\ 0, & x \notin A_i \end{cases}. \quad (1.11)$$

де N – число областей, на яке розбите поле зору X . Множини однакової яскравості A_i попарно не перетинаються і складають все поле зору $X: A_i \cap A_j = \emptyset, i, j = 1, \dots, N; \bigcup_{i=1}^N A_i = X$.

Якщо під зміною умов реєстрації розуміти варіації світлового потоку випромінювання, що висвітлює сцену, то ці зміни позначаються на значеннях яскравості $c_1, \dots, c_N$, але розбиття на області $A_1, \dots, A_N$ залишається незмінним. Таким чином, перетворення, який не міняє структуру зображення, складаються у всіляких перетвореннях яскравості ділянок (множин) $A_1, \dots, A_N$ поля зору X .

Форма зображення f , згідно [3], визначається як множина можливих проєкцій вихідного (ідеального) зображення в $(N+1)$ -мірному лінійному просторі:

$$v_f = \{f(x) = \sum_{i=1}^{N+1} c_i \chi_i(x)\}, \quad x \in X, \quad c_i \in (-\infty, \infty), \quad i = \overline{1, N+1}. \quad (1.12)$$

Поняття форми зображення, згідно (1.12), запроваджується за умови, що поле зору X розбите на множини $A_1, \dots, A_{N+1}$, причому всі точки кожного з цих множин на будь-якому зображенні з v_f , мають однакову яскравість. Якщо ця умова не дотримується, то зображення f складноструктуроване.

У морфологічному аналізі кольорових зображень в якості моделі зображення використовується векторнознакова функція, задана на поле зору. Її значення в кожній точці поля зору визначається спектральними характеристиками об'єкта і спектральним складом освітлення.

Форма кольорового зображення визначається в залежності від розв'язуваної задачі як інваріант перетворень яскравості, кольору або яскравості і кольору вихідного зображення [3].

Зміна зображення об'єкта внаслідок змін умов спостереження можна описати перетворенням $F(\cdot)$, діючим на яскравість пред'явленого зображення $f(\cdot)$:

$$g(x) = F(f(x)) = \sum_{i=1}^N F(c_i) \chi_i(x), x \in X \quad (1.13)$$

де $g(\cdot)$ – зображення об'єкта, отримане з зображення $f(\cdot)$ внаслідок зміни умов спостереження.

В результаті таких перетворень всі точки поля зору зображення $f(\cdot)$, що мають одну і ту ж яскравість c , на зображенні g також матимуть однакову яскравість $F(c)$, в загальному випадку, відмінну від c .

Можливі випадки, коли дві або більше області різної яскравості зображення $f(\cdot)$, на зображенні $g(\cdot)$ матимуть одну і ту ж яскравість.

Тоді виконується умова $F(c_i) = F(c_j)$, $i \neq j$ і множини A_i і A_j – на зображенні $f(\cdot)$ виявляться «склеєними» на зображенні $g(\cdot)$. З огляду на, що складність структури зображення визначається числом N [4], то геометрична форма областей однакової яскравості зображень при «склеюванні» множин спрощується, і можна вважати, що зображення $g(\cdot)$ по формі не складніше, ніж $f(\cdot)$.

1.3 Аналіз методів сегментації складноструктурованих зображень

Сегментація зображень – це процес поділу зображення на кілька сегментів (множина пікселів, також званих 30 суперпікселями). Сегментація зображень зазвичай використовується для того, щоб виділити об'єкти і кордони (лінії, криві, і т.д.) На зображеннях. Більш точно, сегментація зображень – це процес присвоєння міток кожному пікселю зображення, так що пікселі з однаковими мітками мають загальні візуальні характеристики.

Сегментація складноструктурованих зображень до кінця не методизований процес. До методизації процесу сегментації можна виділити два підходи. Перший підхід заснований на акцентуванні перепадів яскравості на зображенні, іншими словами на пошуку «краю» зображення. У це поняття можуть бути включені і межі сегментів, тобто і процес сегментації. Другий підхід до методизації процесу сегментації заснований на визначенні однорідних яркостей (кольорів) або однородностей типу текстур.

При сегментації зображень необхідно ввести деякі базові поняття, які введемо з урахуванням того, що використовується растровий метод синтезу зображень і елементом зображення – множиною є піксель, як атрибути якого виступають координати і яскравість або RGB коди, в разі кольорового зображення. Нехай маємо множина F пікселів, що описують (характеризують) сегмент зображення або зображення в цілому. Точка N_0 називається внутрішньою точкою множини F , якщо всі точки досить малого кола з центром в точці N_0 належать F . Точка N_1 , що належить або яка не належить множині F , називається граничною точкою множини F , якщо будь-яке коло з центром N_1 , містить точки, як належать, так і не належать F .

Множина всіх граничних точок множини F називається кордоном множини F . Точка N називається зовнішньою точкою множини F , якщо разом з цією точкою множини F не належить і деяке коло з центром в N .

Зв'язковою областю називається множина D точок площини xOy , що володіє наступними властивостями:

- D складається з одних внутрішніх точок (властивість відкритості);
- будь-які дві точки, що належать D можна з'єднати безперервної лінією, що цілком складається з точок D (властивість зв'язності).

Множина точок, що складається з області D і її межі, називається замкнутою областю і позначається $\bar{D}$. Межа точки N на площині xOy називається довільна область, яка містить точку N . Під ε околицею кінцевої точки N_0 будемо розуміти коло радіуса ε .

Довільну функцію Z двох змінних, задану на площині xOy , будемо вважати комплексною функцією однієї комплексної змінної, яка визначається в параметричній формі

$$Z = z(t) = x(t) + jy(t), \quad (1.14)$$

де $x(t)$ і $y(t)$ – речові безперервні функції змінної t , $t \in [0, K]$;

K – кількість пікселів, що входять в зображення функції Z на площині xOy .

Якщо $x(t)$ і $y(t)$ неперервні і мають безперервні похідні, причому $(x'(t))^2 + (y'(t))^2 \neq 0$, $t \in [0, K]$, то крива Z називається гладкою. Замкнуту криву, яка не має точок самопересічення, будемо називати замкнутим контуром або просто контуром.

Надалі будемо вважати, що межі розглянутих областей складаються з кінцевого числа замкнутих контурів, розрізів і точок. Число частин, з яких складається весь кордон області D , називається порядком зв'язності цієї області (рис. 1.2).

В інтелектуальних системах, в основу яких покладено аналіз і класифікація зображень, використовують методи сегментації, які можна розділити на два класи: автоматичні – не потребують взаємодії з користувачем і інтерактивні – використовують для користувача введення безпосередньо в процесі роботи.

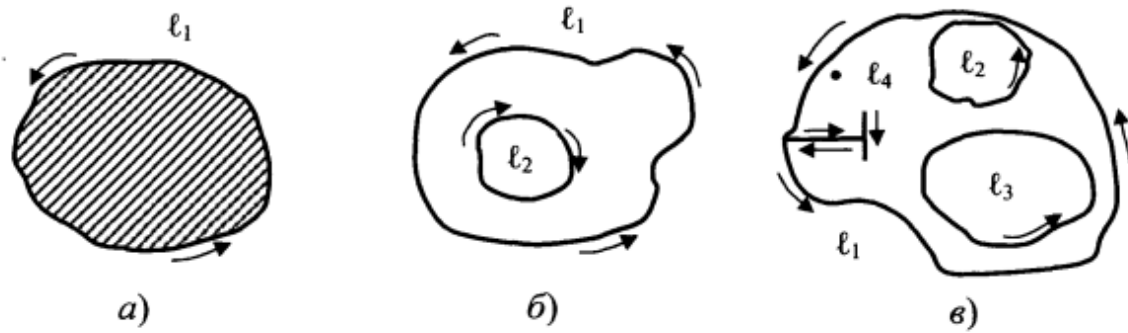


Рисунок 1.2 – Приклади областей з різною зв'язністю:

а) – однозв'язна; б) – двозв'язна; в) – чотиризв'язна

Виділяють два завдання автоматичної сегментації:

- виділення областей зображення з відомими властивостями;
- розбиття зображення на однорідні області.

Між цими двома постановками завдання є принципова різниця.

У першому випадку завдання сегментації полягає в пошуку певних областей, про які є апіорна інформація (наприклад, ми знаємо колір (RGB-код), форму областей, або цікавлять нас області представляють собою зображення відомого об'єкта).

Методи цієї групи вузько спеціалізовані для кожного конкретного завдання. Сегментація в такій постановці використовується в основному в задачах машинного зору (наприклад, аналіз сцен, пошук об'єктів на зображенні). У другому випадку завжди апіорна інформація про властивості областей не використовується, але на саме розбиття зображення накладаються деякі умови (наприклад, всі області мають бути однорідні за кольором і текстурі). Так як при такій постановці завдання сегментації не використовується апіорна інформація про зображених об'єктах, то методи цієї групи універсальні і застосовні до будь-яких зображень.

В основному сегментація в такій постановці застосовується на початковому етапі вирішення задачі, для того щоб отримати уявлення зображення в більш зручному вигляді для подальшої роботи [4].

На результати сегментації досить суттєво впливає як вибір критерію однорідності, так і спосіб побудови однорідних областей. Далеко не завжди для зображення є єдино «правильна» сегментація. У разі сегментації складноструктурованого зображення сегментація не має однозначної відповіді. Тому немає і об'єктивного критерію оцінки якості розбиття складноструктурованого зображення.

Для грубої оцінки якості методу в конкретній задачі зазвичай фіксують кілька властивостей, якими повинна володіти хороша сегментація. Якість роботи методу оцінюється в залежності від того, наскільки отримана сегментація володіє цими властивостями. Найбільш часто використовуються наступні суб'єктивні властивості:

- однорідність регіонів (однорідність кольору або текстури).
- несхожість сусідніх регіонів.
- гладкість межі регіону.
- невелика кількість дрібних «дірок» всередині регіону та т.д.

Для сегментації зображень розроблено кілька універсальних методів і методів. Оскільки спільного рішення для задачі сегментації складноструктурованих зображень не існує, для ефективного вирішення конкретних прикладних задач, що застосовуються методи доводиться модифікувати відповідно до знань предметної області.

1.3.1 Порогова сегментація

Порогові перетворення займають центральне місце в прикладних задачах сегментації зображень. Ряд завдань обробки зображень пов'язаний з перетворенням півтонування (багатоградаційного) в бінарне (двоградаційне) зображення. Перетворення здійснюється для того, щоб скоротити інформаційну надмірність зображення, залишити в ньому тільки ту інформацію, яка потрібна для вирішення конкретного завдання [4].

Бінаризація зображення є обов'язковою процедурою, що передуює затурного аналізу зображень [5].

Очевидний спосіб виділення об'єктів з навколишнього фону полягає у виборі порогового значення T , що розмежовує моди розподілу рівня яскравості пікселів. Тоді будь-який піксель (x, y) , в якому значення яскравості $f(x, y) > T$, називається точкою об'єкта, а в іншому випадку – точкою фону.

За допомогою дворівневого порогового перетворення точка (x, y) класифікується як належить об'єкту одного класу, якщо $f(x, y) > T_2$, якщо якщо $T_1 < f(x, y) < T_2$ – об'єкту іншого класу, і якщо $f(x, y) < T_1$ – фон.

Якщо вибір порогового значення T здійснюється з урахуванням деякої локальної характеристики (наприклад, середньої яскравості в околиці розглянутої точки зображення), то такий поріг називається адаптивним. При цьому вихідне зображення на підставі аналізу рівня фону розбивається на локальні області, в кожній з яких використовується відповідне граничне значення. Завдання сегментації, що вимагають застосування кількох порогових значень, найкраще вирішуються з використанням методів вирощування областей [5].

З практичної точки зору при значних варіаціях яскравості пікселів зображень пороговий метод сегментації по яскравості самотійно застосовується вкрай рідко, а в сукупності з іншими може використовуватися на стадії попередньої обробки зображень. Серед методів, що дозволяють вибрати величину порога, найбільш відомі гістограмного методи.

Суть гістограмного методу полягає в тому, щоб побудувати бімодальне гістограму зображення, вибрати поріг в точці мінімуму і сегментувати зображення з обраним пороговим значенням [6].

Залежно від умов отримання зображення порогове значення може вибиратися одноразово, або адаптивно для кожної локальної області зображення. При цьому гістограма може мати кілька піків для відповідної кількості об'єктів.

Методи з використанням гістограми найбільш ефективні в порівняно з іншими методами сегментації зображень, оскільки при їх реалізації здійснюється лише один прохід по пікселям. Одним з недоліків цих методів є те, що в умовах нерівномірності освітлення або наявності великої кількості об'єктів гістограма яскравості не має бімодальної структури. Тому для застосування гістограмного методу необхідно провести локалізацію об'єкта. Для цих цілей використовують скануючі маски, або розбиття зображень.

Найважливішим плюсом порогових методів є їх обчислювальна ефективність. Тому порогові методи доцільно застосовувати для швидкої, нехай і грубої обробки зображення на початковій фазі сегментації. При цьому найважливішим гідністю гістограмного методу є їх адаптивність до розподілу яскравості об'єкта і фону, яка виражається в можливості адекватного вибору порогового значення для сегментації, причому в різних умовах отримання вихідних даних [7].

Однак така адаптивність в більшості практично значимих випадків носить локальний характер і можлива лише в деякій околиці об'єкта після його локалізації. Для досягнення стійкості сегментації доцільно спочатку адаптуватися до спостережуваного розподілу яскравості зображень, а вже потім проводити сегментацію з оптимальними параметрами [7].

Подальшим розвитком порогових методів є методи вирощування бластей [8]. В основному процедура вирощування областей використовується для отримання окремих регіонів, однак, застосовуючи цю процедуру послідовно або одночасно для декількох регіонів, можна отримати розбиття всього зображення. існують різні стратегії вибору центрів регіонів (seeds) і вирощування регіонів [8].

Метод вирощування областей враховує зв'язність областей і на практиці дозволяє набагато успішніше вирішувати багато завдань, де до цього застосовувалося граничне перетворення. Недоліком методу є зростаюча обчислювальна складність, а основні труднощі становить пошук центрів кристалізації і визначення моменту закінчення росту областей.

1.3.2 Методи сегментації, що засновані на кластеризації

Як атрибути пікселя зображення можна використовувати предствалення кольору в деякому колірному просторі, прикладом метрики (міри близькості) може бути евклідова відстань між векторами в просторі ознак. Найбільш відомі методи кластеризації, які використовуються для сегментації зображень: k -середніх (узагальнений метод Ллойда), ЕМ-метод (англ. Expectation-maximization) [9].

Основна проблема методів кластеризації, полягає в тому, що топологія точок в просторі інформативних ознак або не враховується зовсім, або враховується опосередковано (наприклад, використовуючи координати точки як одна з ознак).

Недоліком методів кластеризації, є те, що вони погано працюють на зашумлених зображеннях: часто губляться окремі точки регіонів, утворюється багато дрібних регіонів, і т.п. [10].

Це особливо характерно для складноструктуріруємих зображень. Модифікацією методів кластеризації є метод зсуву центру мас (метод mean-shift). Основна ідея методу полягає в тому, що центри кластерів відповідають піках щільності розподілу даних [11].

Даний метод має ряд переваг таких, як відсутність необхідності в початкових знаннях про кількість кластерів або про яких-небудь їх ознаках. У рамках методу знаходиться значення функції оцінки щільності. Єдиним параметром методу є розмір вікна h . Поряд з перевагами цього методу існує і ряд недоліків:

- процедура обчислення вектора середнього зсуву вимагає великих обчислювальних ресурсів;
- чутливість до викидів (в результаті сегментації методом mean-shift з'являються невеликі області, що не несуть інформації про семантичної структурі зображення, але вимагають витрат на обробку при інтелектуальному аналізі результатів сегментації);

– чутливість до розмірів вікна, (так як єдиним параметром кластеризації є розмір вікна h , то його значення значно впливає на результат).

У роботі [12] описані спроби усунути дані недоліки, розробивши новий метод агломеративного сегментації на основі методу mean-shift і агломеративного стратегії.

У роботі [13] усунені основні недоліки класичного методу mean-shift. Для скорочення часу пропонується обчислювати вектор середнього зсуву не для всіх точок зображення, або шляхом виконання процедури mean-shift паралельно для кожної точки, що дозволило добитися істотного збільшення швидкості роботи при незначному зниженні якості сегментації.

1.3.3 Методи сегментації, що засновані на градієнтних операторах

Найбільш поширений підхід до сегментації зображень представляє група методів, заснованих на градієнтних операторах. При цьому підході завдання сегментації формулюється як задача пошуку кордонів областей. Методи пошуку меж добре розроблені для напівтонових зображень.

Існує множина підходів до виділення кордонів, але практично все можна розділити на дві категорії: методи, засновані на пошуку максимумів, і методи, засновані на пошуку нулів.

У методах, заснованих на пошуку нулів, визначаються точки перетину вираження другої похідної з віссю абсцис, зазвичай це нулі Лапласіан або точки нульового рівня (zero crossing edge detectors). Як крок передобработки до виділення кордонів практично завжди застосовується згладжування зображення, зазвичай, фільтром Гаусса.

Одним із способів виділення кордонів є методи просторового диференціювання, які засновані на припущенні про те, що граничні точки мають велику величину модуля градієнта функції яскравості.

Один з найбільш ранніх методів з виявлення контурів зображення належить Лоуренсу Робертсу. Модуль градієнта зображення обчислюється за допомогою матриць згортки:

$$G_x = \begin{bmatrix} +1 & 0 \\ 0 & -1 \end{bmatrix} * F, G_y = \begin{bmatrix} 0 & +1 \\ -1 & 0 \end{bmatrix} * F, G = \sqrt{G_x^2 + G_y^2}, \quad (1.15)$$

де F – вихідне зображення;

$*$ – оператор згортки;

G – модуль градієнта зображення.

Оператор Робертса (1.15) є найбільш простим і швидким методом виділення контурів.

Якщо апроксимувати гістограму модуля градієнта рівномірним розподілом ймовірностей, то величина порога буде визначатися як подвоєне середнє значення модуля градієнта [14].

Так як біологічні системи зорового сприйняття, використовують головним чином обриси контурів, а не поділ об'єктів, то для виділення контуру довільного напрямку використовується модуль градієнта поля яскравості, а перепади яскравості представляються модулем «дискретного градієнта» [15].

Часто, з метою підвищення стійкості до шуму, перед застосуванням фільтрації зображення зазвичай розмивають. Завдяки коммутативності оператора Лапласа і фільтра Гауса, можна одночасно здійснювати розмиття і пошук кордонів [16].

Для сегментації сильно зашумлених зображень часто використовується перетворення Хью. Застосовані в ньому, крім градієнтних методів, різноманітні прийоми кластерного аналізу гістограм і методи оцінки, в кінцевому підсумку так перетворюють вихідні дані, що цікавлять характеристики утворюють кластери (скупчення) [17].

Метод вододілу є відносно новим методом сегментації зображень, заснованих на обчисленні градієнта. У цьому методі розглядається абсолютна

величина градієнта зображення як топографічної поверхні. Поняття вододілу засноване на представленні зображення як тривимірної поверхні, з двома просторовими координатами і рівнем яскравості.

Пікселі, що мають найбільшу абсолютну величину градієнта яскравості, відповідають лініям вододілу, які представляють кордону областей.

Безпосереднє застосування методу сегментації по вододілах зазвичай призводить до надлишкової сегментації, викликаній шумом і іншими локальними нерівностями на градієнтному зображенні [18]. Метод вододілу досить широко застосовується в задачах розпізнавання гістологічних зображень [19].

При сегментації складноструктурованих зображень необхідно виходити з того, що завданням сегментації є синтез вирішального правила, яке дозволяє визначити, чи належить досліджуваній піксель до шуканого сегменту.

Як вирішальне правило можуть бути обрані дві послідовні процедури: фільтрація або віконне перетворення і порогова обробка. У разі лінійної фільтрації перша процедура формалізується відповідно до виразу:

$$u_{ij} = \sum_{\gamma=0}^{K-1} \sum_{\beta=0}^{L-1} w_{\gamma\beta} x_{i-\text{int}\left(\frac{K}{2}\right)+\gamma, j-\text{int}\left(\frac{L}{2}\right)+\beta}, \quad (1.16)$$

де $w_{\gamma\beta}$ – вагові коефіцієнти маскує послідовності;

x_y – яскравість пікселя з координатами (i, j) ;

$K \times L$ – розмір фільтруючої маски.

Сегментація виконується, відповідно до виразу:

$$y_{ij} = f(x) = \begin{cases} 1, & \text{якщо } f(u_{ij}) \geq T_{\text{пор}} \\ 0, & \text{якщо } f(u_{ij}) < T_{\text{пор}} \end{cases}, \quad (1.17)$$

де y_{ij} – елемент бінарного зображення шуканого сегмента;

$f(u_{ij})$ – деякий нелінійне перетворення;

T_{nor} – коефіцієнт, який підбирають емпірично або встановлюється на основі розвідувального аналізу [20].

Однак, використання процедури (1.16) при пошуку сегментів дає неоднозначні результати, так як ефективність виділення сегмента залежить не тільки від величини вагових коефіцієнтів, що входять в маску, але і від її структури і розмірності.

У загальному випадку розмір маски повинен відповідати розмірам мінімальних сегментів, що виділяються на зображенні, а її структура повинна адаптуватися до топології кордонів сегмента.

1.3.4 Модельна сегментація

Основне припущення методу сегментації за допомогою моделі ґрунтується на тому, що цікавлять структури або органи мають повторювані геометричні форми. Отже, можна знайти вірогідну модель для пояснення змін форми об'єкта і потім, сегментуючи зображення, накладати обмеження, використовуючи цю модель як апіорну.

Таке завдання включає в себе:

- приведення тестованих зображень до стандартного положення;
- розподіл усіх уявлення змін наведених зразків;
- статистичний висновок для моделі і зображення.

Сучасні методи сегментації, засновані на апіорних знаннях, містять активні моделі форми і зовнішності, активні контури, деформуються шаблони і методи встановлення рівня.

Недоліком цього методу є те, що вимірювання зображення, відповідні окремої області, можуть мати значну зміну внаслідок таких ефектів, як відображення, текстура поверхні, похибки пристроїв реєстрації та інші.

Завдання текстурної сегментації сформульована на основі марковських моделей зображення і полягає в пошуку прихованого марковського поля з найбільшою апостеріорної ймовірністю [21].

Це завдання зводиться до оптимізаційних завдань надзвичайну складність, за винятком деяких окремих випадків, коли, наприклад, Марківського поле є авторегресійну або гаусовим.

Для вирішення загальної задачі запропоновано метод модельованого відпалу, який успішно застосовувався для практичного вирішення інших завдань.

Відомо, однак, що модельований отжиг не визначає однозначно метод відшукування глобального оптимуму, а лише вказує, що шуканий метод належить великому класу методів, що відрізняються один від одного процесом зміни спеціального параметра, так званої «температури» відпалу.

Тільки при певній залежності температури від часу методи модельованого відпалу забезпечують збіжність до глобального максимуму оптимізується функції.

Вибір залежності температури від часу становив основну проблему при використанні рекомендацій дослідників при вирішенні практичних завдань [21]. На сьогоднішній день, як моделі, дуже часто використовуються гомеостатичні моделі або моделі м'яких обчислень. При сегментації складноструктурованих зображень як такі моделі використовуються штучні нейронні мережі. Для реалізації процедури виділення контуру за допомогою нейронної мережі було прийнято таку модель [22]. На вхід нейронної мережі подається зображення вікна спостереження розміром 21×21 .

Це вікно задовольняє наступним вимогам. Вікно завжди містить елемент контуру, тобто в будь-якому представленому вікні міститься межа перепаду яркостей об'єкта і фону. На зображенні може бути присутнім адитивний гаусів

шум, який перешкоджає правильній роботі методу. На виході повинні бути отримані основні параметри прямої, що апроксимує кордон перепаду яскравості. На зображенні може бути присутнім адитивний гаусів шум, який перешкоджає правильній роботі методу.

Для реалізації цієї моделі була обрана нейронна мережа типу багатошаровий перцептрон. Мережа містить два шари.

У першому шарі 10 нейронів, в другому – 3.

Усі штучні нейрони використовували логістичну функцію активації, архітектура мережі зображена на рисунку 1.3.

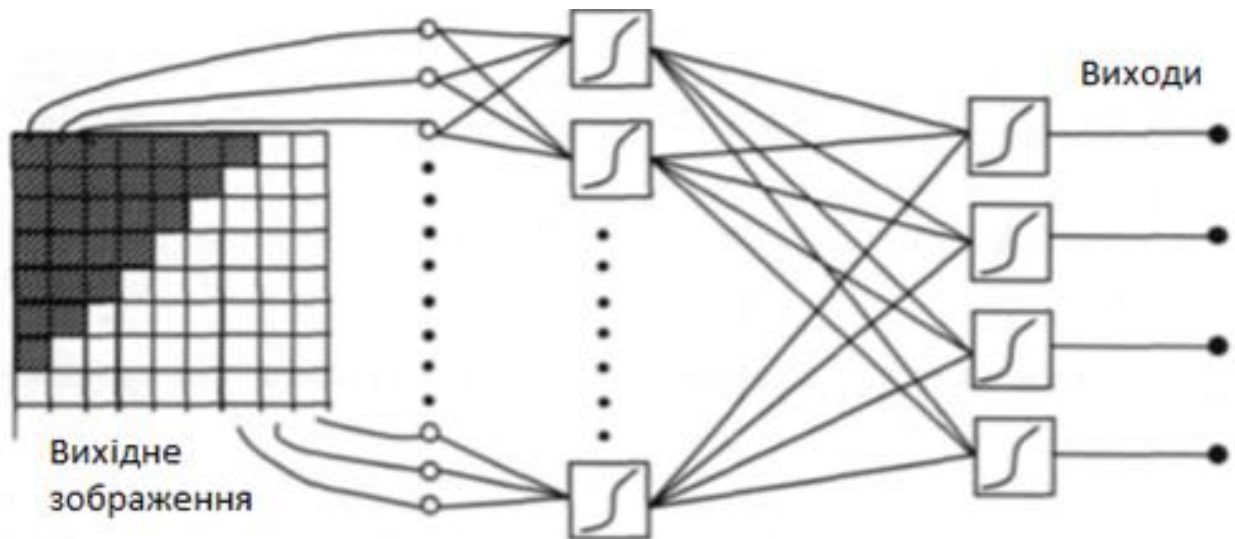


Рисунок 1.3 – Архітектура нейронної мережі,
що реалізує процедуру виділення контуру

При навчанні, яскравість кожного елемента зображення приводилась до діапазону $[0,1; 0,9]$, і вихідне зображення у вигляді вектора 1×441 подавалося на вхід мережі.

При цьому в якості цільового вектора приймався вектор параметрів, кожен елемент якого також приведений до діапазону $[0,1; 0,9]$.

При використанні штучних нейронних мереж операції сегментація і класифікація можуть бути проведені одночасно. При цьому виникає

необхідність в проведенні попередньої обробки зображень до подачі їх на вхід нейронної мережі.

Це обумовлено тим, що застосування нейронних мереж з безпосередньою подачею на їх вхід пікселів вихідного зображення пов'язано з наступними проблемами:

- при пред'явленні пікселів зображення на вхід нейронної мережі виникають складності при розпізнаванні з урахуванням кутів повороту ділянок текстур щодо країв зображення;

- так як текстура – просторове властивість, то вона повинна оброблятися в межах певної області – вікна певного розміру. З використанням такого вікна виникає проблема вибору його розміру.

При виборі цього розміру малим можуть спостерігатися значні помилки усередині ділянок з однорідною текстурою, так як при цьому результат обробки стає статистично нестійким. Якщо ж вибрати таке вікно великого розміру, виникає значна кількість помилок розпізнавання поблизу кордонів текстурних ділянок за рахунок захоплення вікном двох і більше різних текстур.

Перша проблема може бути вирішена шляхом застосування попередньої статистичної обробки зображення. При цьому розраховуються статистики, які не залежать від повороту текстур, всередині аналізованого вікна. Як статистики можуть бути застосовані статистики відмінності рівнів сірого [23] і статистики просторової залежності рівнів сірого [24], або статистики мікротекстури Лоза [24].

Однак застосування двох перших статистик більш прийнятно, так як на відміну від статистик, запропонованих [25], вони враховують взаємне просторове розташування пікселів один щодо одного.

Для вирішення другої проблеми вибирається вікно невеликого розміру, і в цьому вікні по вихідного зображення розраховуються простору статистик. При цьому забезпечується досить низький рівень помилок, що виникають за рахунок змішування статистик поблизу кордонів двох або більше ділянок з

різними текстурами, що відбувається через захоплення цих текстурних ділянок вікном, в якому проводиться аналіз.

Усунення статистичної нестійкості отриманих даних, обумовленою малим розміром вікна при розрахунку текстурних властивостей, проводиться шляхом фільтрації отриманих просторів текстурних статистик за допомогою фільтра, що згладжує, що зберігає контури, з використанням вікна більшого розміру [26].

Таким чином, при існуванні стабільних відмінностей в яскравості окремих областей поля зору застосовуються порогові методи. При наявності стійкої зв'язності всередині окремих сегментів найбільш ефективні методи нарощування областей.

Ідея цих методів полягає в тому, що відбувається угруповання сусідніх елементів зображення з однаковими або близькими рівнями яскравості, а потім об'єднання їх в однорідні області.

Один з методів – центроїдне зв'язування – припускає вибір стартових точок або за допомогою оператора, або автоматично. Ефективним є метод вододілів, заснований на пошуку локальних мінімумів з подальшою угрупованням навколо них областей по зв'язності.

Метод виділення кордонів доцільно застосовувати в тому випадку, коли над ділянкою досить чіткі і стабільні. Виділення контурних ліній на зображенні найбільш часто використовується в системах технічного зору, засноване на обліку зміни яскравості і надалі її порівнянні з граничним значенням.

Перераховані методи застосовуються для виділення сегментів за критерієм однорідних яскравостей.

Всі методи дуже прийнятні з точки зору обчислювальних витрат, однак, для кожного з них характерна неоднозначність розмітки отворів в реальних ситуаціях через необхідність застосування евристик (вибір порогових значень збіги яркостей, вибір цифрових масок і т.д.). Заслуговує на увагу, в зв'язку з

цим, метод багатозначною розмітки, заснований на комбінації різних прийомів для зниження невизначеності.

Важливе практичне значення мають методи, що допускають паралельну обробку для прискорення процесу розмітки на основі логічного аналізу сусідніх елементів.

Для опису та сегментації властивостей зображень, а саме, однорідності, шорсткості, регулярності, застосовують текстурні методи, діляться умовно на дві категорії: статистичні та структурні. Прикладом статистичного підходу є використання матриць збігів, що формуються шляхом вихідних зображень, з подальшим підрахунком статистичних моментів та ентропії. При структурному підході, наприклад, на основі мозаїки Вороного, будується множина багатокутників. Багатокутники із загальними властивостями об'єднують в області. Для дослідження загальних властивостей часто використовують ознаки моменти багатокутників. Однак, після сегментації часто виникають перешкоди у вигляді, як розрізнених змін ізольованих елементів зображення, так і у вигляді спотворень деяких зв'язкових областей.

1.4 Опис процедур та методів класифікації зображень

На сьогоднішній день існують різні підходи для аналізу і класифікації складноструктуріруємих зображень. Класичний підхід, застосовуваний для аналізу і класифікації зображень, може бути представлений у вигляді послідовності, що складається з чотирьох етапів: отримання зображення, його попередня обробка з можливим подальшим виділенням об'єктів, отримання вектора ознак і класифікація зображень на їх основі [27].

Основним елементом будь-якого завдання розпізнавання зображень є відповідь на питання: чи належать дані (вхідні) зображення до класу зображень, який представляє даний еталон? Здавалося б, відповідь можна отримати, порівнюючи безпосередньо зображення з еталонами (або їх вектори

ознак). Однак, виникає ряд труднощів і проблем, специфічних, особливо, при створенні систем технічного зору (СТЗ):

- зображення пред'являються на складному тлі;
- зображення еталона і вхідні зображення відрізняються становищем в поле зору;
- вхідні зображення не збігаються з еталонами за рахунок випадкових перешкод;
- відмінності вхідних і еталонних зображень виникає за рахунок зміни освітленості, підсвічування, локальних перешкод;
- еталони і зображення можуть відрізняти геометричні перетворення, включаючи такі складні як афінніє і проектні.

Для вирішення завдання в цілому і на окремих її етапах застосовуються різні методи сегментації, нормалізації і розпізнавання.

Основні методи обробки і розпізнавання системою технічного зору (СТЗ) зображень можна класифікувати відповідно до схеми, наведеної на рисунку 1.4.

Для вирішення цього завдання слід перевірити рівність нулю норми.

На схемі вказані основні процедури і методи обробки від початкового етапу сприйняття поля зору за допомогою датчиків, наприклад, телекамери, до кінцевого, яким є розпізнавання.

Процес розпізнавання зображень є складною багатоетапною процедурою. Багатоетапність (ієрархічність) обумовлена тим, що різні завдання обробки насправді тісно пов'язані і якість вирішення однієї з них впливає на вибір методу вирішення інших.

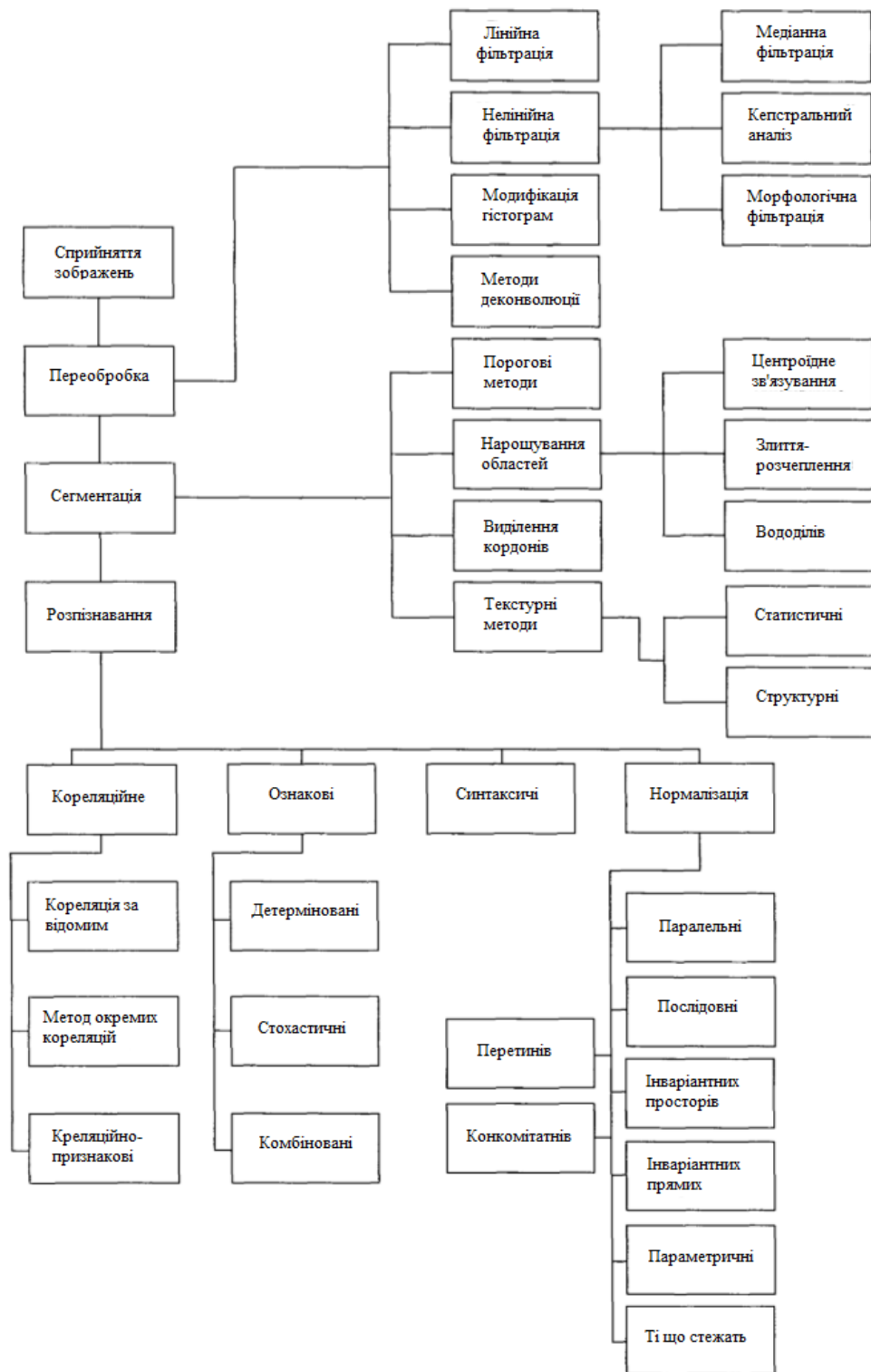


Рисунок 1.4 – Основні процедури і методи розпізнавання зображень

Так вибір методу розпізнавання залежить від конкретних умов пред'явлення вхідних зображень, в тому числі характеру фону, перешкод, інших зображень, і пов'язаний з вибором методів попередньої обробки, сегментації, фільтрації [28].

Для класифікації об'єктів за формою їх зображень формулюється наступна постановка задачі. Передбачається заданим набір форм зображень.

До різних об'єктів у вигляді множин $v_k, k = \overline{1, K}$, кожне з яких є множиною зображень k -го об'єкта при різних умовах реєстрації. Передбачається, що їм відповідають проектори $P, k = \overline{1, K}$.

Завдання полягає у виборі того класу зображень, до якого належить пред'явлене для класифікації зображення g одного з об'єктів. Для вирішення цього завдання слід перевірити рівність нулю норми $\|g - P_k g\| = 0$.

Однак, в реальних умовах в силу неточності моделі, наявності шумів і т.п. пред'явлення зображення може не належати жодному з заданих класів $v_k, k = \overline{1, K}$, тоді необхідно знайти той клас, до якого найближче за формою зображення g . Якщо в якості міри близькості вибирається квадрат норми $\|g - P_k g\|^2$, то задача класифікації зводиться до задачі мінімізації:

$$\|g - P_k g\|^2 \Rightarrow \min_{k \in (1, \dots, K)} . \quad (1.20)$$

Цей мінімум дорівнює нулю при повній відповідності пред'явленого зображення моделі формування. Якщо мінімум в (1.20) досягається при декількох індексах $k_{i_1}, \dots, k_{i_m}$, то g класифікується як зображення одного з об'єктів з номерами $k_{i_1}, \dots, k_{i_m}$.

1.5 Постановка задачі дослідження

Головною метою дослідження є виявлення ефективних гібридних методів для класифікації складноструктурованих зображень. Для виконання необхідно проаналізувати різноманітні методи сегментації, визначити особливості роботи з бінарними зображеннями.

Також однією з основних задач стане дослідження мережових та нейромережових структур для задач класифікації, а також визначення з методичним та програмним забезпеченням, що необхідні для формування вирішальних модулів.

Завершальним етапом дослідження має стати використання гібридних методів для класифікації:

- визначення вхідних даних;
- тестування методів;
- аналіз результатів.

2 ІНТЕЛЕКТУАЛЬНІ ОПЕРАТОРИ СЕГМЕНТАЦІЇ ТА МЕТОДОЛОГІЯ СПЕКТРАЛЬНОГО ОПИСУ КРИВИХ СКЛАДНОЇ ФОРМИ СКЛАДНОСТРУКТУРОВАНИХ ЗОБРАЖЕНЬ

2.1 Узагальнений метод сегментації складноструктурованих зображень

Під сегментацією в загальному випадку розуміють процедуру подання зображення у вигляді багатозв'язної області. Границі областей описуються бінарним зображенням – замкнутим контуром. Процес сегментації передбачає декомпозицію вихідного зображення на множина зображень або для отримання початкового зображення відповідної багатозв'язної області, що відбиває результат сегментації.

Більшість методів сегментації включає послідовну обробку зображення градієнтним і пороговим операторами. Використовують два анізотропних градієнтних оператора, в результаті чого отримують градієнтне зображення, кожен піксель якого характеризується двома параметрами, обумовленими як

$$\begin{aligned} G I_x &= dG/dx, \\ G I_y &= dG/dy, \end{aligned} \tag{2.1}$$

де G – вихідне зображення;

$G I$ – зображення, отримане обробкою градієнтним оператором;

x і y – напрями, що характеризують анізотропію градієнтного оператора.

Багато зображень не мають чіткого морфологічного опису сегментів і відносяться до класу зображень. При цьому складність структуризації обумовлена як природою самого зображення, так і зашумленістю зображення. Дуже часто зображення отримують в умовах різної освітленості, як всього зображення, так і його частин.

Також виникають ситуації, коли одні і ті ж сегменти на різних зображеннях мають різну колірну забарвлення. У багатьох випадках складноструктуровані зображення виходять з добре структурованих зображень за рахунок того, що об'ємні зображення з явно вираженими сегментами проєктуються на двовимірну площину.

Проекції сегментів накладаються один на одного, при цьому утворюються псевдо сегменти або спотворені сегменти, що призводить до зниження цінності апріорних даних

2.2 Синтез морфологічного оператора для препарування зображень

Під морфологічним оператором розуміємо кортеж арифметичних і логічних операцій, здійснюваних над множинною пікселів зображення $G1$, що потрапили в прямокутне «вікно» W , в центр якого поміщений активний піксель. Під активним пікселем розуміємо поточний піксель зображення, що переглядається шляхом порядкової розгортки. Отже, процес препарування включає дві процедури:

- процедура формування «вікна» морфологічного оператора;
- процедура операцій над пікселями, що входять в «вікно» морфологічного оператора.

Вікно W морфологічного оператора отримуємо в результаті накладення «вікон», що мають форму вектора.

$$W=W_0 \cup W_1 \cup W_2 \cup W_3, \quad (2.2)$$

де W_Θ – «вікно» у вигляді вектора, орієнтованого в просторі під кутом Θ до горизонтальної осі декартових координат, $\Theta=\overline{0,3}$.

На рисунку 2.1 представлено «вікно» W , складене з пятиелементних «вікон» W_Θ .

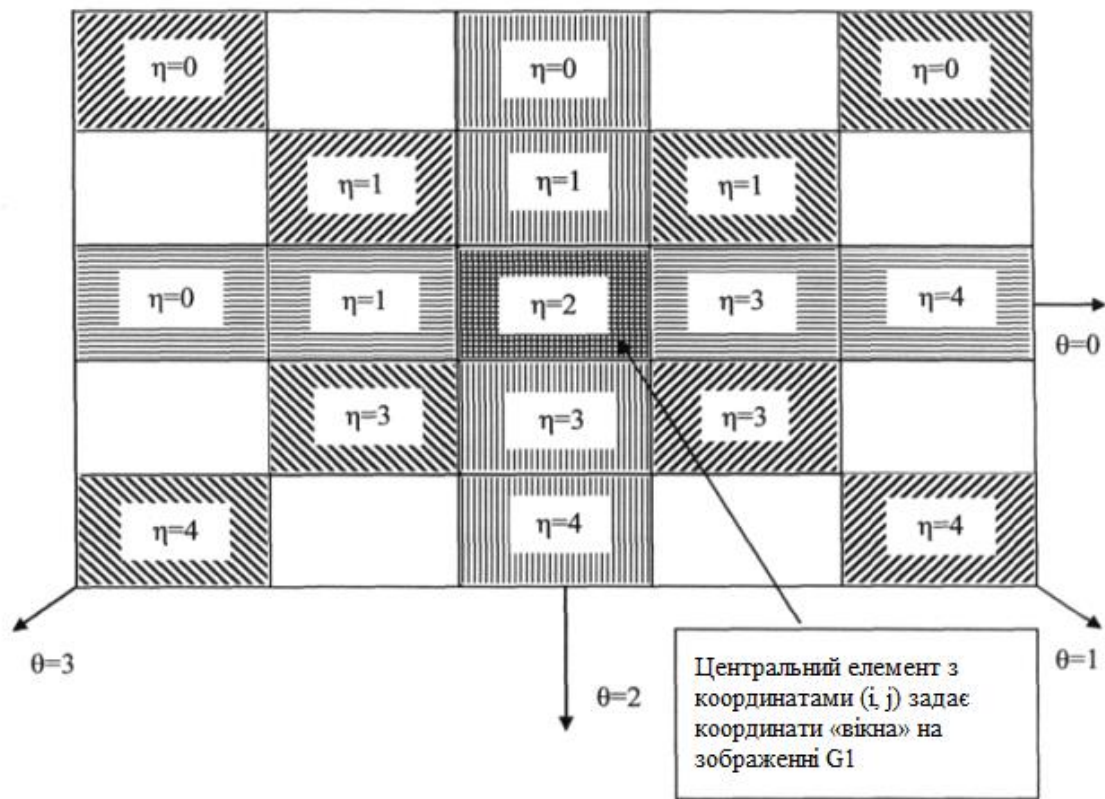


Рисунок 2.1 – Структура вікна для морфологічного оператора

Кожне п'ятиелементне «вікно» орієнтоване певним чином щодо центрального елемента з координатами (i, j) . При цьому центральний елемент є загальним для всіх чотирьох п'ятиелементної «вікон». Орієнтацію п'ятиелементного «вікна» W_θ щодо центрального елемента визначає параметр θ . Порядковий номер пікселя в п'ятиелементній «вікні» задає параметр η .

Таким чином, в п'ятиелементному «вікні» – вікно з орієнтації $\theta = 0$ входить множина наступних пікселів:

$$W_0 = \{w_{00}, w_{01}, w_{02}, w_{03}, w_{04}\}. \quad (2.3)$$

При цьому, щоб визначити координати пікселів на зображенні $G1$ відповідні координатам «вікна» W_θ , необхідно знати координати «вікна» W , тобто координати центрального елемента (i, j) .

Координати пікселів (θ, η) щодо «вікна» W морфологічного оператора, визначаються за допомогою таблиці 2.1.

У таблиці 2.1 представлені координати пікселів, щодо активного елементу вікна, за допомогою яких кожному елементу «вікна» ставиться у відповідність елемент зображення $G1$. «Вікно» морфологічного оператора має координати активного пікселя (i, j) . Для морфологічного оператора з квадратним «вікном» розміром 5×5 пікселів координати (θ, η) представлені в таблиці 2.1.

Таблиця 2.1 – Таблиця відповідності локальних координат пікселів η у вікнах W_θ з глобальними координатами

$\theta \backslash \eta$	0	1	2	3	4
1	$i,j-2$	$i,j-1$	i,j	$i,j+1$	$i,j+2$
2	$i-2,j-2$	$i-1,j-1$	i,j	$i+1,j+1$	$i+2,j+2$
3	$i-2,j$	$i-1,j$	i,j	$i+1,j$	$i+2,j$
4	$i-2,j+2$	$i-1,j+1$	i,j	$i+1,j-1$	$i+2,j-2$

Схема методу перетворення пікселя зображення $G1$ в піксель зображення $G2$ представлена на рисунку 2.2.

Метод реалізується в результаті послідовного виконання морфологічних операцій в чотирьох «вікнах» W_θ . В якості вхідної інформації використовується масив ψI .

На початку, організовується цикл по θ . Цикл дозволяє здійснити морфологічні операції в чотирьох п'ятиелементних «вікнах».

У кожному п'ятиелементному «вікні» може бути реалізовано одна з двох морфологічних операцій: «дилатація» або «ерозія».

Якщо в «вікні» в здійснюється операція «дилатація», то

$$\forall w_{\theta\eta} \in W_{\bar{\theta}} \left[g2_{i_{\bar{\theta},\eta}} = \max \{ G1(l_{\bar{\theta},\eta}) \} \right]. \quad (2.4)$$

Аналогічно, якщо для «вікна» в реалізується морфологічна операція «ерозія», то

$$\forall w_{\theta\eta} \in W_{\bar{\theta}} \left[g2_{i_{\bar{\theta},\eta}} = 0 \right]. \quad (2.5)$$

Кожному $\bar{\theta} \in \{\theta\}$ відповідає певний напрям кордону сегмента.

Всього обрано чотири напрямки: захід-схід; північний схід; північ-південь; північний захід. З огляду на те, що масив градієнтів яскравості $G1$ є векторним, кожне з чотирьох напрямків ділиться на пряме і на зустрічне, тобто відхиляється на 180 від прямого напрямку.

У цьому випадку $(-180^\circ < \psi l_{\theta,\eta} < 180^\circ)$.

Морфологічна операція, яка виконується в п'ятиелементному «вікні», визначається за наступним вирішальним правилом

$$\begin{aligned} &\text{ЯКЩО } \forall w_{\theta\eta} \in W_{\bar{\theta}} \left[\left(\left| \psi_{\bar{\theta},\eta} \right| > \bar{\theta} \cdot \pi/4 - \pi/8 \right) \text{ and } \left(\left| \psi_{\bar{\theta},\eta} \right| > \bar{\theta} \cdot \pi/4 + \pi/8 \right) \right] \\ &\text{ТО у вікні } W_{\bar{\theta}} \text{ виконується операція «дилатація», інакше «ерозія»}. \end{aligned} \quad (2.6)$$

З огляду на, що кожне значення параметра $\theta = \bar{\theta}$ ($\theta = \overline{0,3}$) дає два напрямку, що відрізняються на 180 , продукційних правил типу (2.5) має бути теж два.

Друге продукційне правило запишеться так

$$\begin{aligned} &\text{ЯКЩО } \forall w_{\theta\eta} \in W_{\bar{\theta}} \left[\left(\left| \psi_{\bar{\theta},\eta} \right| > \pi + \bar{\theta} \cdot \frac{\pi}{4} - \frac{\pi}{8} \right) \text{ and } \left(\left| \psi_{\bar{\theta},\eta} \right| > \pi + \bar{\theta} \cdot \frac{\pi}{4} + \frac{\pi}{8} \right) \right] \\ &\text{ТО у вікні } W_{\bar{\theta}} \text{ виконується операція «дилатація», інакше «ерозія»}. \end{aligned} \quad (2.7)$$

Продукційні правила (2.5) і (2.7) дозволяють збільшити яскравість пікселів, що потрапили в «вікно», за умови, що у всіх цих пікселів фаза лежить всередині сектора, що задається правою частиною продукції.

Продукції (2.5) і (2.7) можуть бути записані у вигляді однієї продукції шляхом введення нової змінної C , яка приймає два значення $C=\{0,1\}$:

ЯКЩО

$$\forall w_{\theta\eta} \in W_{\bar{\theta}} \left[\left(\left| \psi_{\bar{\theta}\eta} \right| > C \cdot \pi + \bar{\theta} \cdot \frac{\pi}{4} - \frac{\pi}{8} \right) \text{ and } \left(\left| \psi_{\bar{\theta},\eta} \right| > C \cdot \pi + \bar{\theta} \cdot \frac{\pi}{4} + \frac{\pi}{8} \right) \right]$$

ТО у вікні Wg виконується операція «дилатація», інакше «ерозія». (2.8)

Продукція (2.8) для одного активного пікселя з координатами (i,j) виконується в блоках 5–11, представлених на рисунку 2.2.

Операцією «дилатації» або «ерозії», процес виконання морфологічного оператора в «вікні» W не закінчується. Якщо в підмножині пікселів «вікна» $W_{\bar{\theta}}$ здійснюється операція «ерозія», то інших подальших операцій над цим підмножиною не проводиться, здійснюється перехід до іншого «вікна» в блоці 3.

При виконанні операції «дилатація» необхідно враховувати, що операція «дилатація» здійснюється тільки в одному з чотирьох підмножин W . Якщо в результаті обчислення продукції (2.8) операція «дилатація» виконується в декількох або у всіх підмножествах множини пікселів в «вікні» W , необхідно вибрати тільки одне підмножина.

Решта пікселі підмножин множини пікселів, що потрапили в «вікно» W заповнюються нулями (за винятком центрального пікселя, який належить всім підмножини Θ).

Для виділення єдиного підмножини $W_{\bar{\theta}}$ з множини W_{Θ} , в якому здійснюється операція «дилатація», у всіх підмножествах з операцією «дилатація» визначається модуль вектора $V_{\bar{\theta}}$ за наступним співвідношенням:

$$|V_{\bar{\theta}}| = \sqrt{(\sum_{\eta=0}^4 g1_{\bar{\theta}\eta}^x)^2 + (\sum_{\eta=0}^4 g1_{\bar{\theta}\eta}^y)^2}. \quad (2.9)$$

Операція «дилатація» здійснюється в том «вікні» $\bar{w}$, в якому (2.9) максимальне, тобто

$$\bar{\theta} = \underset{\bar{\theta}}{\operatorname{argmax}} V_{\bar{\theta}}, \quad (2.10)$$

де $\bar{\theta}$ – підмножина напрямки θ , в яких здійснюється операція «дилатація».

Морфологічний оператор «ерозія» виконується відомими методами морфологічної обробки чорно-білих зображень, описаних, наприклад, в [28].

Виконання морфологічного оператора «дилатація» відрізняється від відомих способів в зв'язку з тим, що в межах «вікна» W він може виконуватися кілька разів з подальшим вибором тільки одного напрямку і виконанням у всіх інших напрямках «вікна» W , в яких виконувалася операція «дилатація», операції «ерозія».

Таку операцію дилатації назвемо «дилатація $\bar{\theta} \rightarrow I$ ». У результаті виконання методів, схеми яких представлені на рисунку 2.3 і рисунку 2.4, синтезується зображення $G2$.

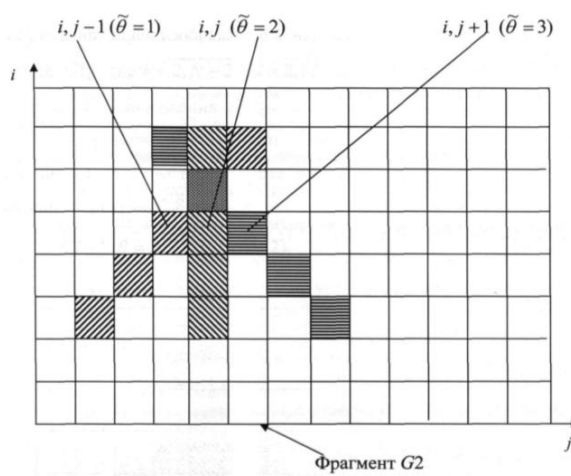


Рисунок 2.3 – Ілюстрація синтезу зображення $G2$

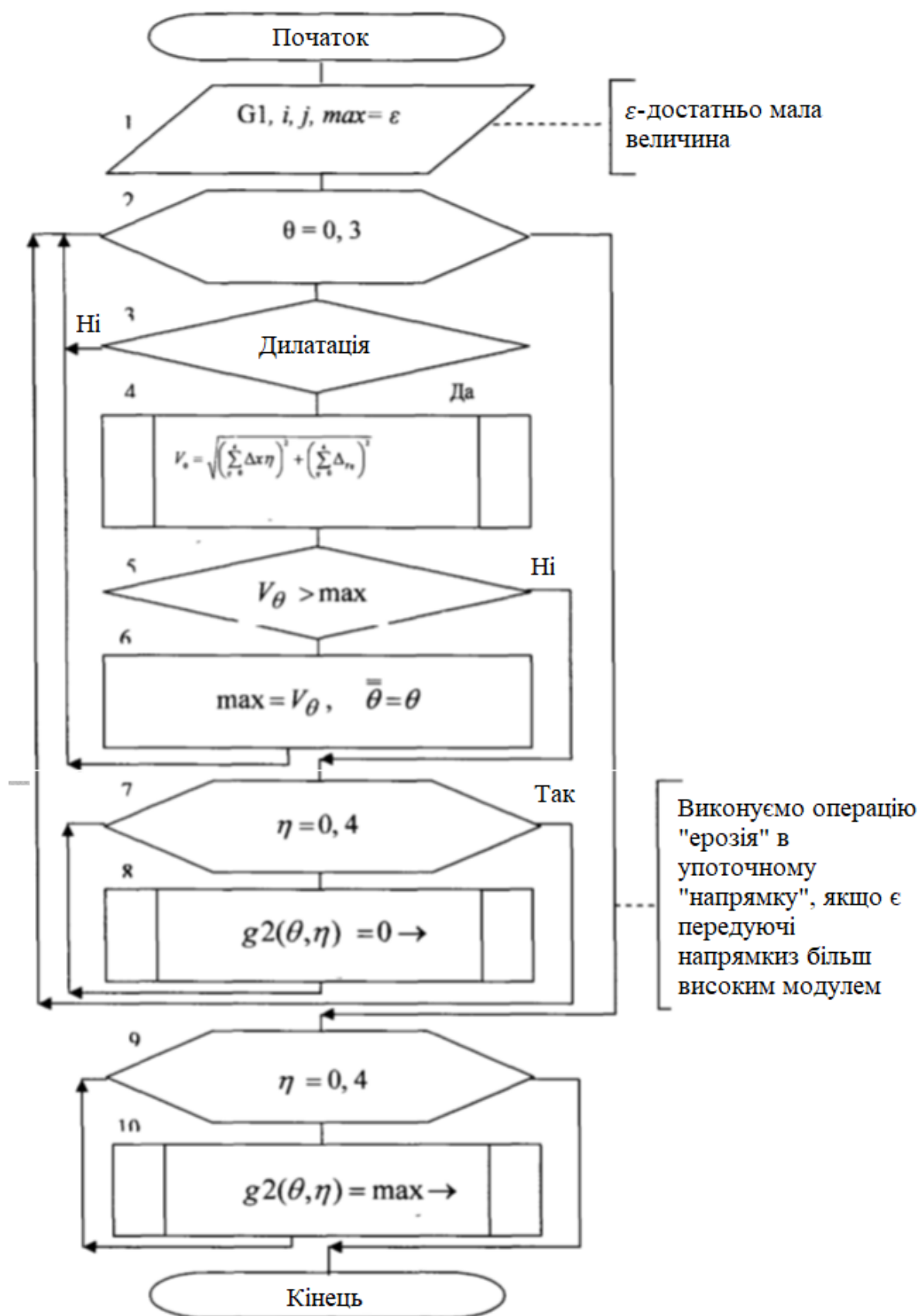


Рисунок 2.4 – Схема методу реалізації морфологічної операції

«дилатація $\theta \rightarrow 1$ »

Синтез зображення $G2$ полягає в тому, що на канву, розміром $(N \times M)$ послідовно наносяться п'ятиелементні маски з центральним елементом, координати якого збігаються з поточними координатами канви (i, j) . При цьому координати центрального елемента п'ятиелементної маски (i, j) , скануючої зображення канви $G2$, збігаються з координатами зображення $G1_{kl}$, де $k = \overline{2, N-2}$; $l = \overline{2, M-2}$.

В процесі сканування канви на нїї накладаються фрагменти, отримані в результаті операції «дилатація $\Theta \rightarrow I$ ». Можливі розташування їх на канві в результаті сканування трьох пікселів канви представлені на рисунку 2.4.

Схема методу синтезу зображення $G2$ з фрагментів, отриманих в результаті операцій «дилатація $\Theta + I$ » над зображенням $G2$ представлена на рисунку 2.5. Процес перенесення пікселів підмножини, що відноситься до напрямку $\overline{\Theta}$, на канву зображення $G2$ здійснюють блоки 6 і 7 методу, представленого на рисунку 2.5.

Так як в процесі перенесення пікселі підмножин переносяться при різних координатах (i, j) (рис. 2.5, блок 5) можуть збігатися, то в блоці 7 рисунка 2.4 здійснюється рекуррентне накопичення.

Це рекуррентне накопичення здійснюється, наприклад, в пікселі з координатами $(i-l, j+l)$ (рис. 2.4). Так як при отриманні зображення $G2$ використовуються необроблені пікселі зображення $G1$, то воно, так само як і зображення $G1$ – градієнтне, отже, при рекуррентном накопиченні використовується формула векторного підсумовування:

$$g2_x^{i,j}(l_{\overline{\Theta},\eta}) = g2_{xi}^{i,j}(l_{\overline{\Theta},\eta}) + g2_x^{i,j}(l_{\overline{\Theta},\eta}), \quad (2.11)$$

$$g2_y^{i,j}(l_{\overline{\Theta},\eta}) = g2_{yi}^{i,j}(l_{\overline{\Theta},\eta}) + g1_y^{i,j}(l_{\overline{\Theta},\eta}), \quad (2.12)$$

де (i, j) – поточні координати «вікна», $\eta = \overline{0, 4}$.

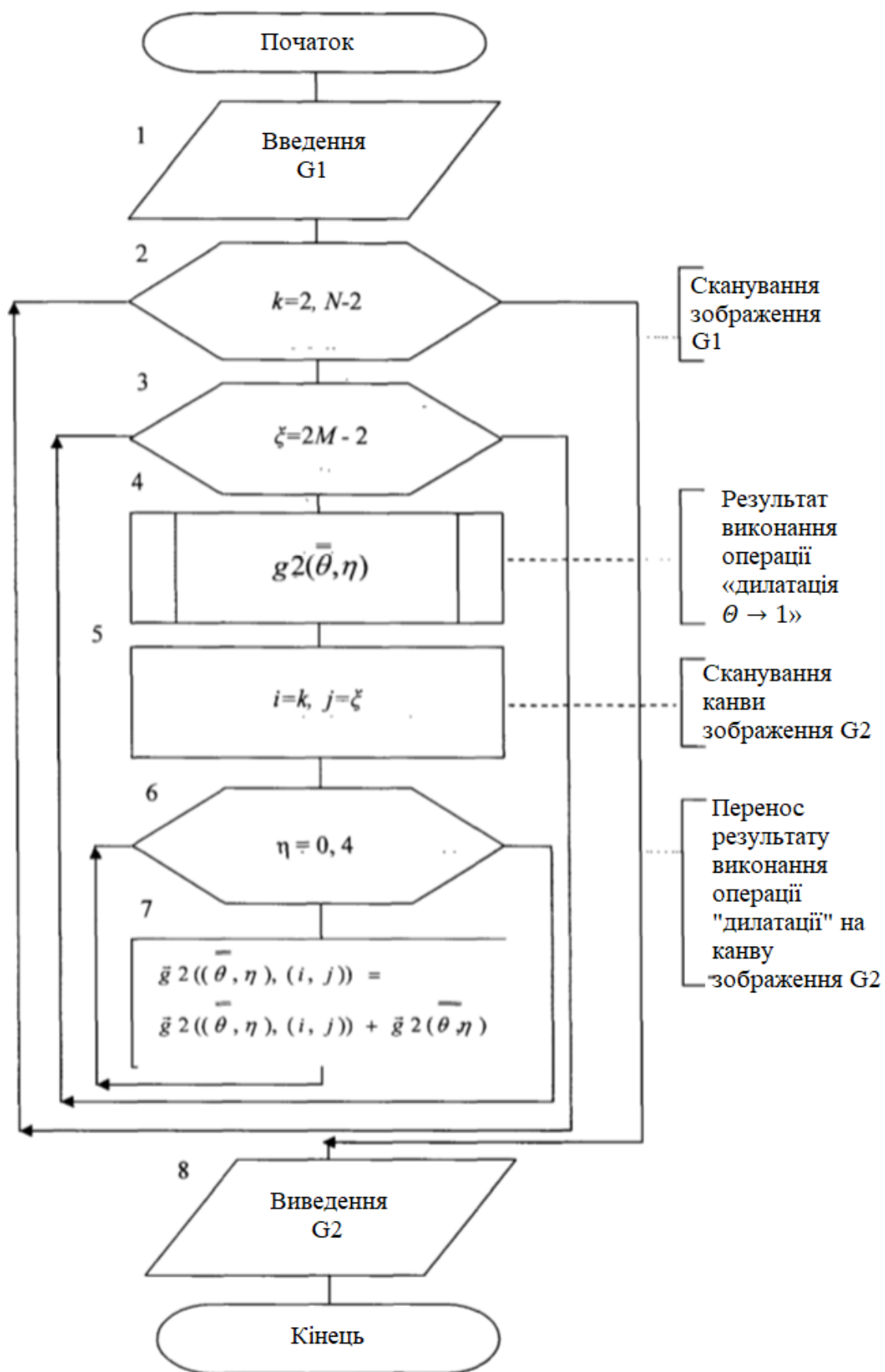


Рисунок 2.5 – Схема методу синтезу зображення

2.3 Особливості обробки бінарних зображень

2.3.1 Нечітка модель формування бінарного зображення границі сегменту

Зображення G_3 отримують з зображення G_2 . В принципі, бінарне зображення меж сегментів G_3 може бути отримана безпосередньо з градієнтного зображення G_1 , тобто без препарування G_1 [28]. При переході до G_3 градієнтне зображення перетвориться у скалярний зображення з бінарними пікселями. Мережева модель перетворення зображення G_2 зображення G_3 показана на рисунку 2.6.

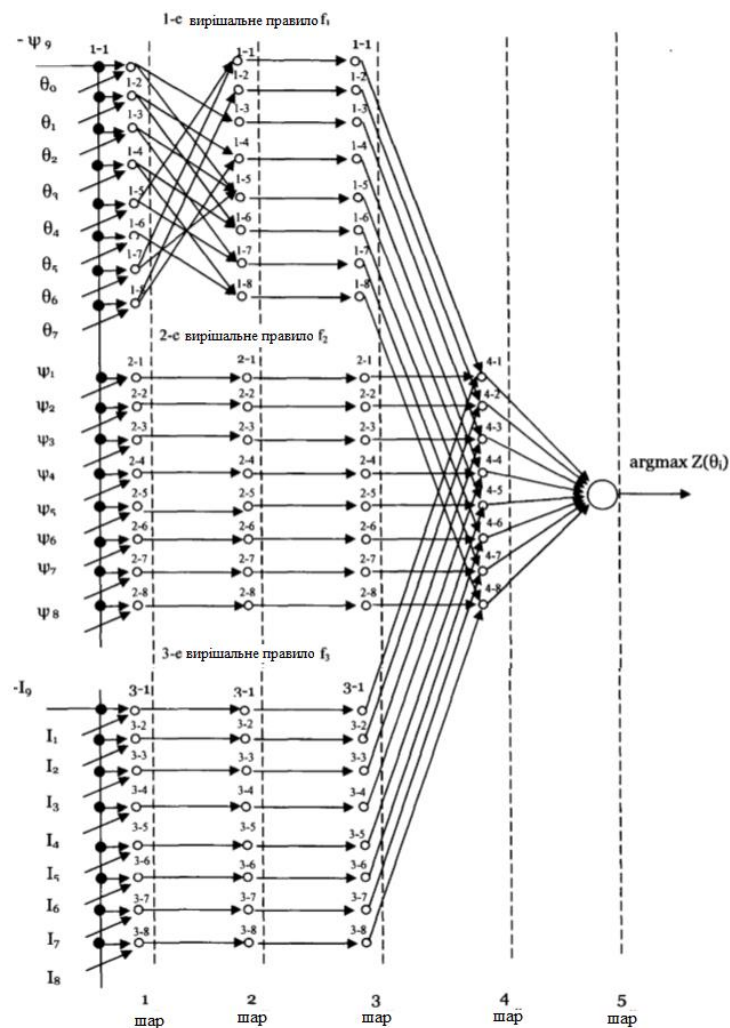


Рисунок 2.6 – П'ятишарова модель нечіткої сітки, що реалізує інтелектуальний оператор виділенням кордонів сегменту

Ця модель включає п'ять шарів. За основу формування шару взята дев'ятиелементная маска з вісьмома напрямками $\theta = \overline{0,7}$, представлена на рисунку 2.7 а). На цьому рисунку піксель, позначений цифрою 9, є вихідним або попереднім пікселем формованої межі сегмента. Щодо дев'ятого елемента можуть бути обрані вісім альтернативних рішень щодо вибору наступного пікселя межі сегмента, які представлені на рисунку 2.7 б) і позначені $\theta_0, \dots, \theta_7$.

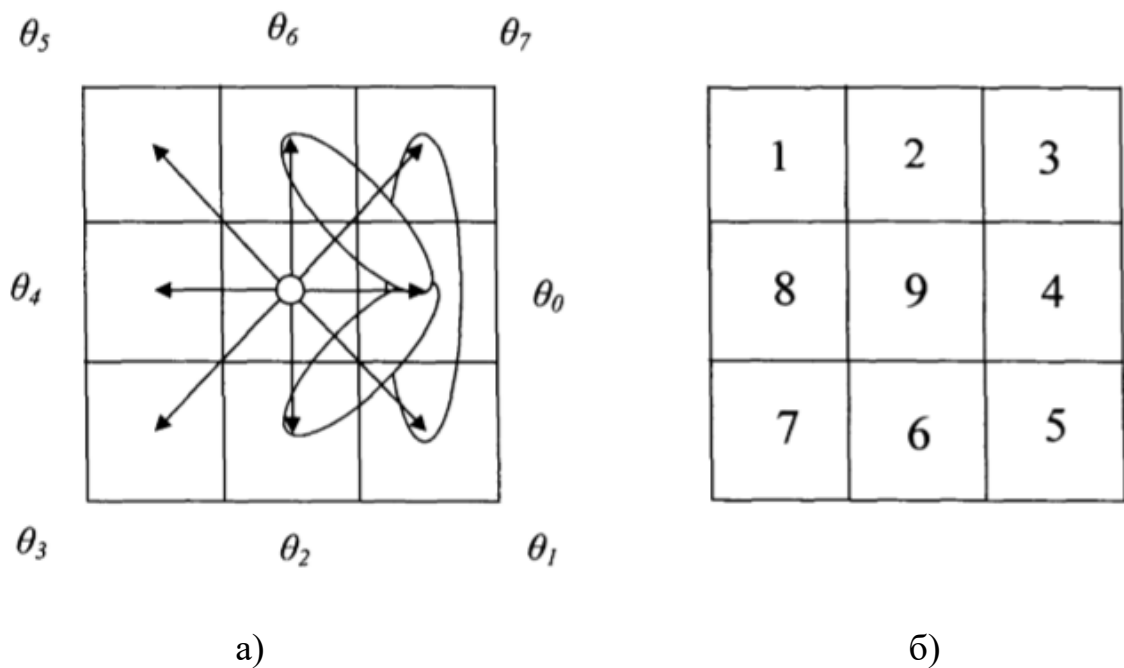


Рисунок 2.7 – Схема:

- а) формування напрямків градієнтів у дев'ятиелементному вікні;
- б) визначення пікселів у дев'ятиелементному вікні

Для вибору відповідної альтернативи можемо сформулювати три вирішальні правила.

На рисунку 2.7 б) ці можливі рухи межі сегмента показані для напрямів градієнтів $\theta_0, \theta_1, \theta_7$ у вигляді трьох еліпсів, великі осі яких перпендикулярні виділеним напрямкам, а менші осі збігаються з можливими градієнтами в дев'ятому пікселі.

Згідно з першим вирішальним правилом обирається той еліпс руху, мала вісь якого найбільш близька до градієнту в пікселі 9, тобто до ψ_9 .

Таким чином, у першому шарі для першого вирішального правила необхідно провести порівняння ψ_9 з усіма θ_i ($i = \overline{0,7}$) і вибрати два з восьми можливих пікселів, які знаходяться на більшій осі еліпса, відповідного напрямку θ_i .

Друге вирішальне правило підтримує твердження про те, що найбільшу впевненість активації має той піксель, градієнт якого найбільш близький до градієнту поточного активізованого пікселя, тобто необхідно перевірити умову:

$$|\psi_9 - \psi_i| \rightarrow \min, \quad i = \overline{1,8}. \quad (2.13)$$

Третє вирішальне правило ґрунтується на тому положенні, що яскравість активується пікселя повинна бути близька до яскравості передус пікселя межі сегмента

$$|l_9 - l_i| \rightarrow \min, \quad i = \overline{1,8}. \quad (2.14)$$

Таким чином, у першому шарі мережевої моделі (рис. 2.10) відбувається формування умов для трьох використовуваних вирішальних правил. Другий шар мережі (рис. 2.6) реалізує математичну обробку умов, отриманих у першому шарі.

Для першого вирішального правила (вузли 1–1 ... 1–8) математична обробка зводиться до знаходження мінімального значення сигналу, що приходить у вузол. У вузол надходять два сигнали, кожен з яких обчислюється в першому шарі як

$$u_i = \psi_9 + \theta_i, \quad i = \overline{1,8}. \quad (2.15)$$

Таким чином, у вузли 1–1 ... 1–8 другого шару приходять два сигнали, що визначаються відповідно до рівняння (2.14).

У кожному з вузлів обчислюється мінімальний сигнал, який надходить на третій шар (вузли 1–1 ... 1–8).

У вузлах 2–1 ... 2–8 і 3–1 ... 3–8 другого шару обчислюються ліві частини рівнянь (2.13) і (2.14), відповідно.

В третьому шарі здійснюється формування нечітких вирішальних правил. Кожне з трьох вирішальних правил здійснює нелінійне перетворення величин, обчислених у вузлах другого шару, у відповідне число, що лежить в діапазоні $0 \dots v$.

Величина v визначає вагу даного вирішального правила, $v \in [0,1]$.

Вага вирішального правила будемо задавати емпірично, з подальшим її уточненням в залежності від виду досліджуваного об'єкта і якості зображень (наявність шуму).

Як нелінійне перетворення використовуємо найпростіші нелінійні перетворення кусково-лінійного типу. Таким чином, маємо три вирішальних правил:

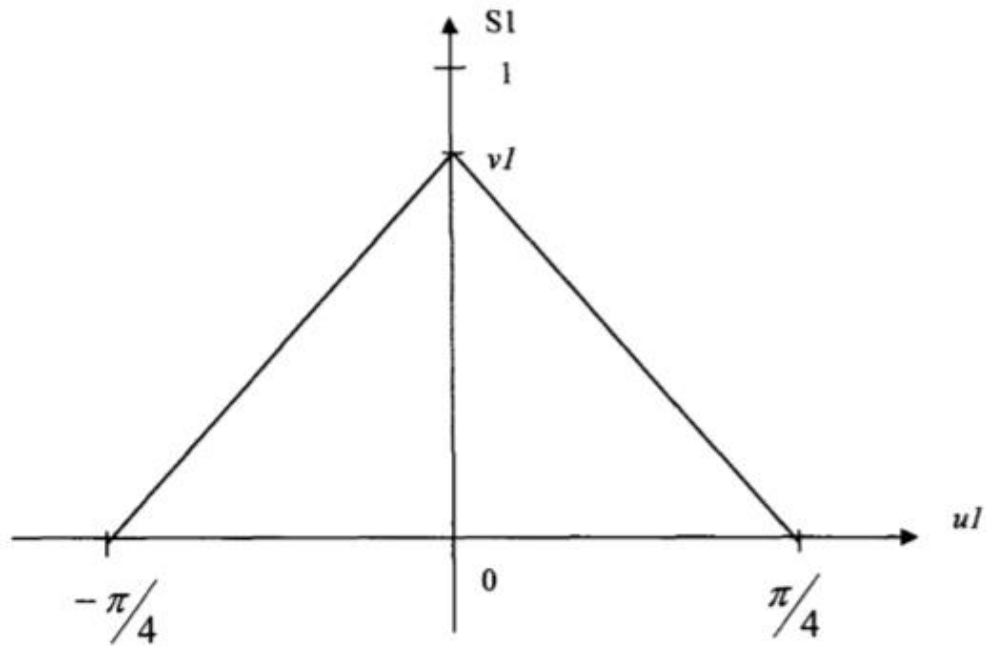
$$S_1 = f_1(u_{1i}); \quad (2.16)$$

$$S_2 = f_2(u_{2i}); \quad (2.17)$$

$$S_3 = f_3(u_{3i}); \quad (2.18)$$

де u_{1i}, u_{2i}, u_{3i} – виходи вузлів другого шару.

Вид нелінійного перетворення $S1$ показаний на рисунку 2.8.

Рисунок 2.8 – Нелінійне перетворення $S1$

Нелінійне перетворення $S1$ описується наступним рівнянням

$$S1 = \begin{cases} 0, & \text{при } u1 < -\pi/4 \\ \left(4 \cdot \frac{v1}{\pi}\right) \cdot u1 + v1, & \text{при } -\pi/4 \leq u1 < 0 \\ -\left(4 \cdot \frac{v1}{\pi}\right) \cdot u1 + v1, & \text{при } 0 \leq u1 \leq \pi/4 \\ 0, & \text{при } u1 > \pi/4 \end{cases} \quad (2.19)$$

Нелінійне перетворення $S2$ дозволяє оцінити впевненість у виборі пікселя межі сегмента на основі порівняння градієнта в поточному і вибираемом пікселі. Вид цього перетворення показано на рисунку 2.9.

Кусково-лінійна апроксимація функції $S2$ представлена рівнянням

$$S1 = \begin{cases} 0, & \text{при } u2 < 0 \\ -\left(2 \cdot \frac{v2}{\pi}\right) \cdot u2 + v2, & \text{при } 0 \leq u2 < \pi/2 \\ 0, & \text{при } u2 > \pi/2 \end{cases} \quad (2.20)$$

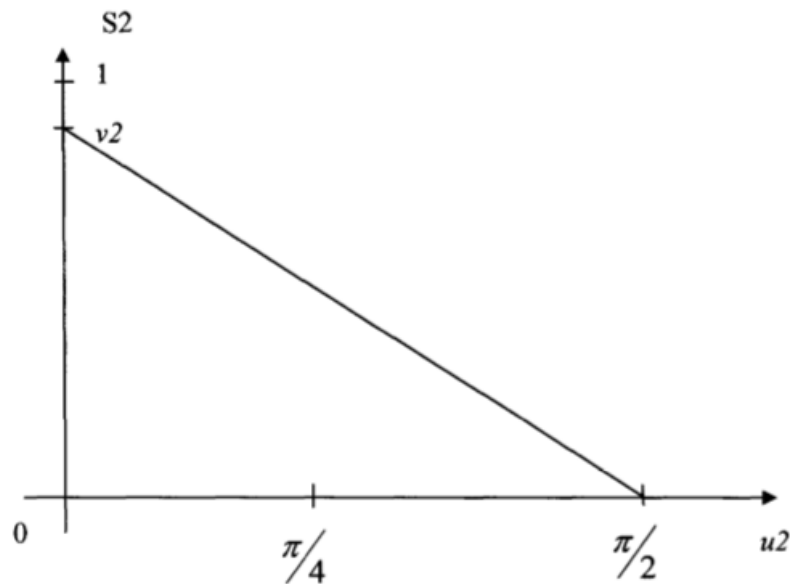


Рисунок 2.9 – Нелінійне перетворення S2

Нелінійне перетворення S3 дозволяє оцінити впевненість у виборі кордону пікселя на основі того, наскільки його яскравість близька до яскравості поточного пікселя.

Вид цього перетворення показаний на рисунку 2.10.

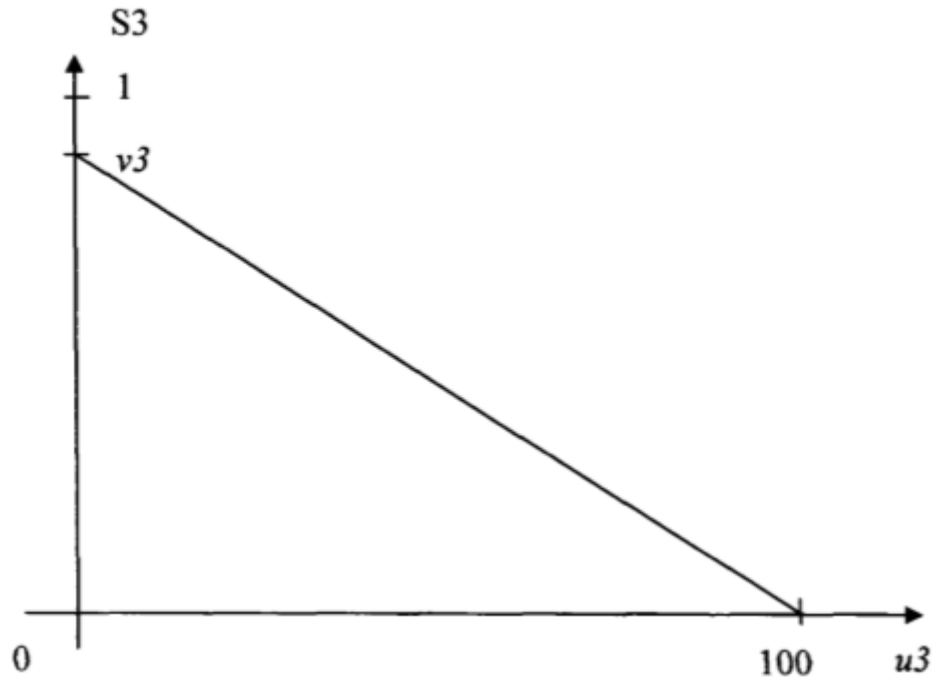
Аргумент нелінійного перетворення, визначається як

$$U3_i = \frac{abs(I_i - I_9)}{I_9 + \varepsilon} \cdot 100, \quad (2.21)$$

де I_9 – яскравість поточного пікселя;

I_i – яскравість пікселів в масці ($i = \overline{1,8}$);

ε – досить мала величина, що визначається шумовими характеристиками зображення (при 255 рівневному квантуванні ε може бути обрана рівною одиниці).

Рисунок 2.10 – Нелінійне перетворення S_3

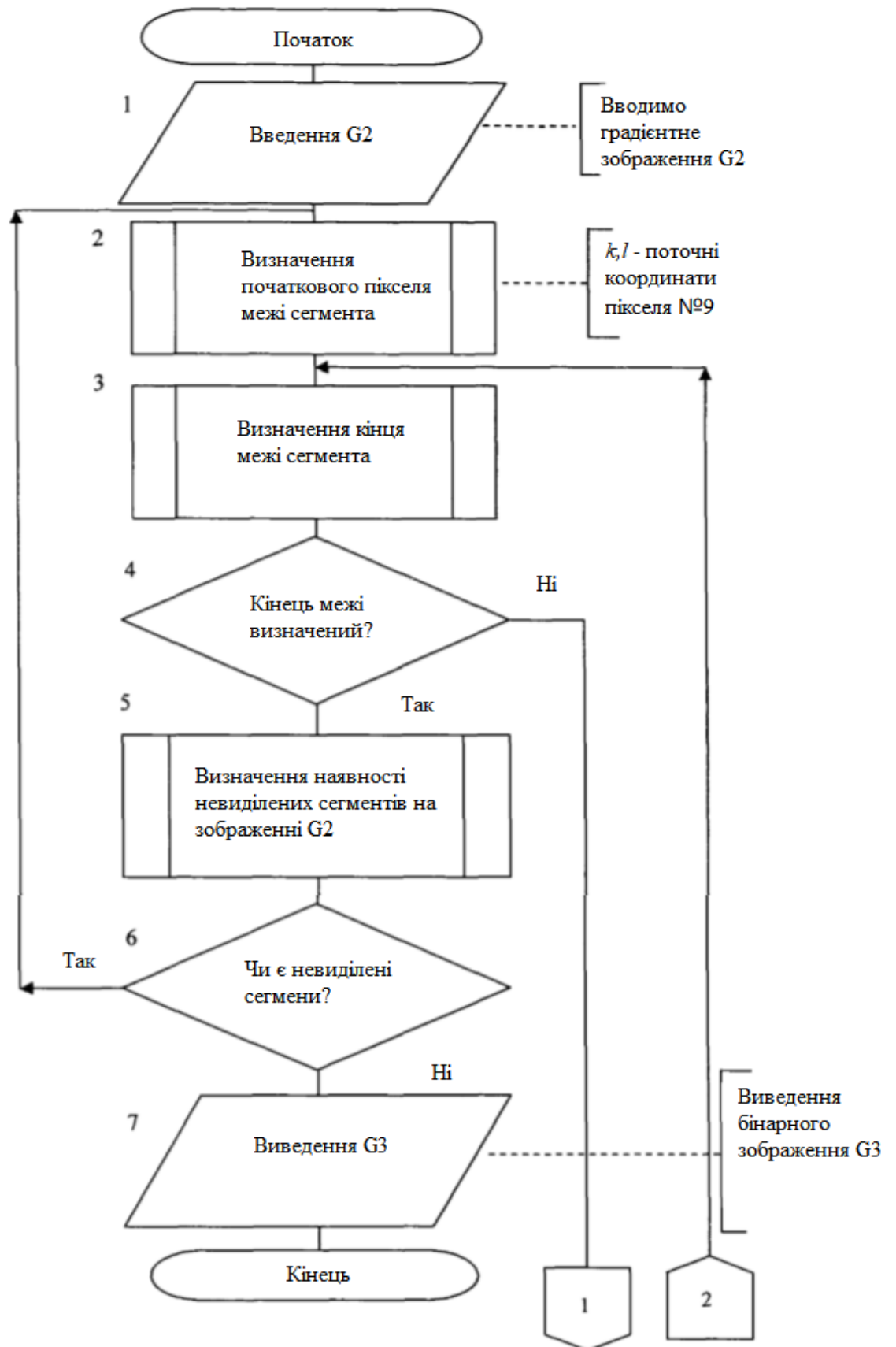
Нелінійна залежність S_3 має наступну кусково-лінійну апроксимацію

$$S_3 = \begin{cases} 0, & \text{при } u_3 < 0; \\ -0.01 \cdot v_3 \cdot u_3 + v_3, & \text{при } 0 \leq u_3 \leq 100; \\ 0, & \text{при } u_3 > 100 \end{cases} \quad (2.22)$$

У четвертому шарі мережі здійснюється агрегація вирішальних правил S_1 , S_2 , S_3 для кожного пікселя 1 ... 8 маски (рис. 2.11). У п'ятому шарі вибирається один з восьми пікселів, агрегируюча функція Z для якого має максимальне значення.

Схема методу, що реалізує нечітку мережу рисунку 2.6, представлена на рисунку 2.11.

Для визначення максимального значення цієї функції необхідно порівняти її значення у восьми точках. Для здійснення порівняння необхідно ввести відстань d в просторі (S_1, S_2, S_3) .



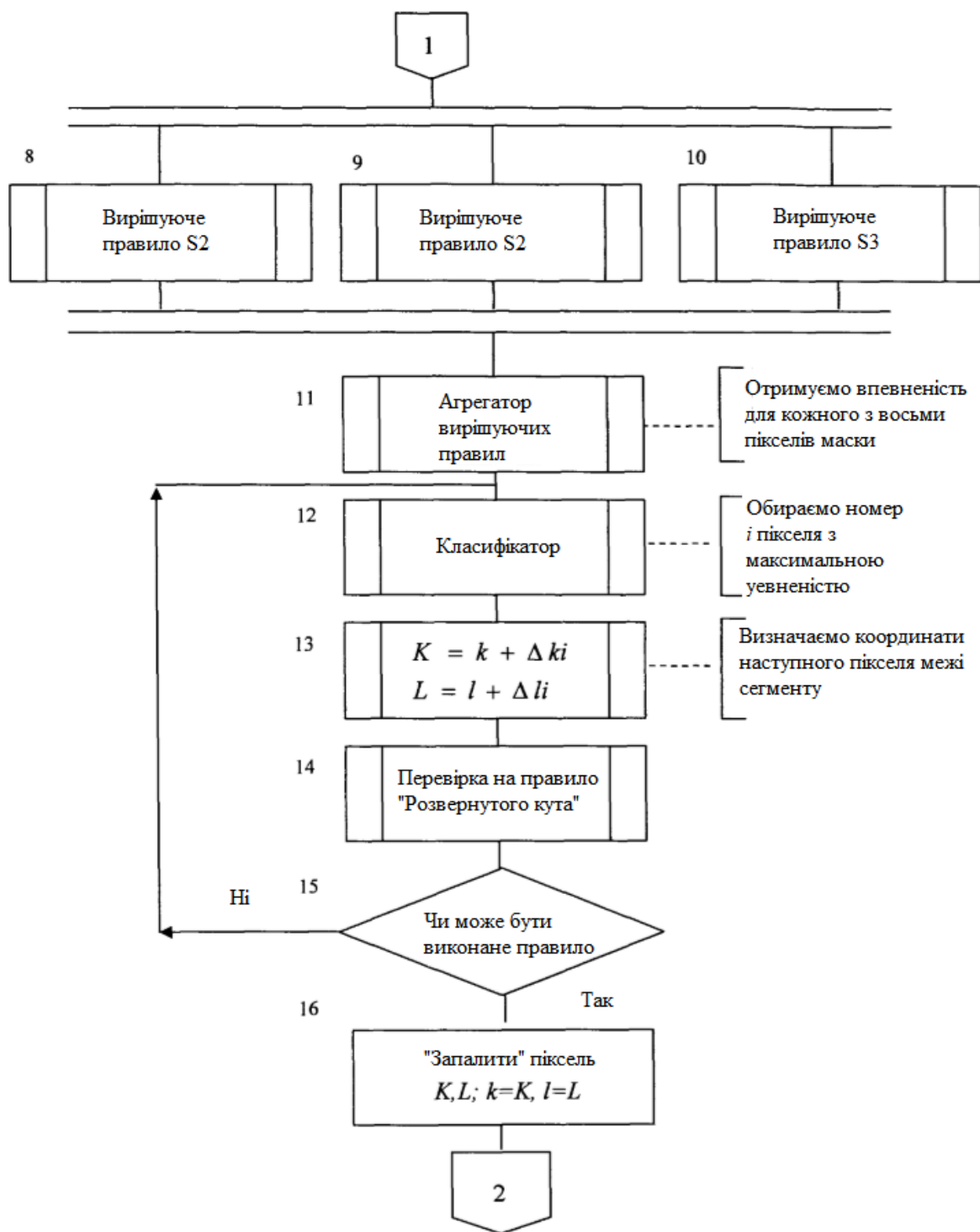


Рисунок 2.11 – Схема методу переходу від градієнтного зображення до бінарного

Це пояснюється тим, що для визначення Z необхідно виконати деяку алгебраїчну операцію над значеннями функцій $S1$, $S2$, $S3$, а для виконання цієї операції необхідно нормувати простір $(S1, S2, S3)$.

Якщо норму ввести через манхетенську відстань, то агрегатор визначається як

$$Z = SI + S2 + S3. \quad (2.23)$$

Процес отримання бінарного зображення $G3$ починається з введення в пам'ять комп'ютера градієнтного зображення $G2$. На зображенні $G2$ визначається початковий піксель межі першого виділеного сегмента. В цей піксель поміщається маска (рис. 2.11), так що дев'ятий елемент маски відповідає координатам до k, l зображення $G2$.

Як критерій вибору початку межі сегмента може служити яскравість елемента векторного масиву $G2$, яка визначається як

$$I = \sqrt{g_x^2 + g_y^2}. \quad (2.24)$$

При виборі початкового пікселя межі сегмента перевіряємо дві умови:

- чи належить піксель, який претендує на початок межі сегмента, «вікна» бінаризованого пікселя (на зображенні $G3$);
- чи перевершує яскравість вибраного пікселя пороговий рівень.

Схема методу, що реалізує блок 2, представлена на рисунку 2.12.

У блоці 2 схеми методу рисунка 2.12 здійснюється перехід від зображення $G2$ до зображення $\tilde{G}2$ згідно співвідношеннями

$$\tilde{g}_{21} = \sqrt{g_{21}^2 + g_{22}^2}, \quad (2.25)$$

$$\tilde{g}_{22} = \arctg g_{22}/g_{21}. \quad (2.26)$$



Рисунок 2.12 – Схема методу пошуку першого пікселя кордону сегмента

Після отримання векторного двовимірного масиву $G2$ формується векторний одновимірний масив R , що отримується в результаті упорядкування по спадаючій яскравості (2.25) масиву $\tilde{G}2$.

Перевірка виконання першої умови здійснюється в блоці 8 (рис. 2.12).

З цією метою вибирається прямокутна маска $W1$, центральний елемент якої встановлюється в піксель зображення $G3$ з координатами (k, l) .

Якщо в масці $W1$ на зображенні $G3$ знаходиться хоча б один бінарізований піксель, то піксель з координатами (k, l) , не може бути початком кордону сегмента. Відзначимо, що ніщо не заважає йому бути пікселем, включеним в кордон сегмента.

Перевірка виконання другого (порогового) умови здійснюється у блоці 10 (рис. 2.12).

Поріг вибирається таким чином, щоб шум або зміна освітленості не впливали на результати сегментації. Для визначення цього рівня необхідно провести статистичний аналіз складової зображення $\tilde{G}2_1$.

З цією метою досліджується гістограма $\tilde{G}2_1$, і аналізується її модальність. Блок 8 (рис. 2.12), в якому здійснюється цей процес, може бути реалізований в автоматичному або інтерактивному режимі.

У блоці 12 (рис. 2.11) визначається наступний піксель кордону сегмента щодо напрямків $\theta_0, \dots, \theta_7$, а в блоці 13 – його координати.

Ці координати стають поточними координатами межі сегменту за умови виконання правила «розгорнутого кута». Це правило вимагає, щоб напрямок руху поточного пікселя відрізнялося від напрямку руху попереднього пікселя не більше ніж на 180° . Так як піксель може рухатися тільки в фіксованих напрямках $\theta_0, \dots, \theta_7$ то дозволені переміщення для цих восьми напрямків відображені в таблиці 2.2.

Таблиця 2.2 – Відповідності дозовлених напрямків руху кордону сегмента

Попередній напрямок	Поточні дозовлені напрямки				
θ_0	θ_0	θ_1	θ_2	θ_6	θ_7
θ_1	θ_0	θ_1	θ_2	θ_3	θ_7
θ_2	θ_0	θ_1	θ_2	θ_3	θ_7
θ_3	θ_1	θ_2	θ_3	θ_4	θ_5
θ_4	θ_2	θ_3	θ_4	θ_5	θ_6
θ_5	θ_3	θ_4	θ_5	θ_6	θ_7
θ_6	θ_4	θ_5	θ_6	θ_7	θ_0
θ_7	θ_5	θ_6	θ_7	θ_0	θ_1

Якщо правило «розгорнутого кута» виконано, то в блоці 14 визначаємо координати наступного пікселя як

$$K = k + \Delta k_i, \quad (2.27)$$

$$L = I + \Delta l_i, \quad (2.28)$$

де i – номер подальшого пікселя кордону сегмента щодо маски, представленої на рисунку 2.7 б);

Δk_i і Δl_i – визначаються згідно з таблицею 2.3.

Таблиця 2.3 – Переклад номера пікселя в прирощення координат

i	1	2	3	4	5	6	7	8
Δk_i	-1	-1	-1	0	1	1	1	0
Δl_i	-1	0	1	1	1	0	-1	-1

Однією з основних процедур виділення кордону сегмента є визначення кінця кордону сегмента.

Як детектор кінця кордону сегмента використовуємо наступне вирішальне правило:

Якщо $\{(подальша\ координата\ пікселя\ кордону\ вийшла\ за\ кадр\ зображення)\}$ або $\{(подальша\ координата\ пікселя\ кордону\ сегмента\ вже\ вибрана)\}$, то «кінець кордону сегмента», інакше «поточний піксель належить кордоні сегмента».

2.3.2 Метод морфологічної обробки бінарних зображень

Розглянемо задачу математичного уявлення складної кривої. Вважаємо, що процедура сегментації завершена, і кожен сегмент представлений довільній кривій (чорно-білим зображенням)

Постановка завдання морфологічного опису кривої полягає в наступному. Нехай ми маємо на кадрі R^2 бінарного зображення деяку множину I кривих

$$\{F_i(x, y)\} \in R, i \in I. \quad (2.29)$$

Припустимо, що крива $F_i(x, y) \in R$ задана векторної функцією $\vec{r}_i = \vec{r}_i(s)$, $0 \leq s \leq S_i$, де змінна s – довжина дуги кривої S_i , тоді похідна векторної функції

$$\frac{d\vec{r}_i}{ds} = \vec{\tau}_i = (\cos \alpha, \cos \beta) \quad (2.30)$$

є множиною одиничних векторів, спрямованих уздовж дотичній до даної кривої (рис. 2.13).

У наведений вище формулі α й β – кути між дотичній і позитивними напрямками осей Ox і Oy , відповідно.

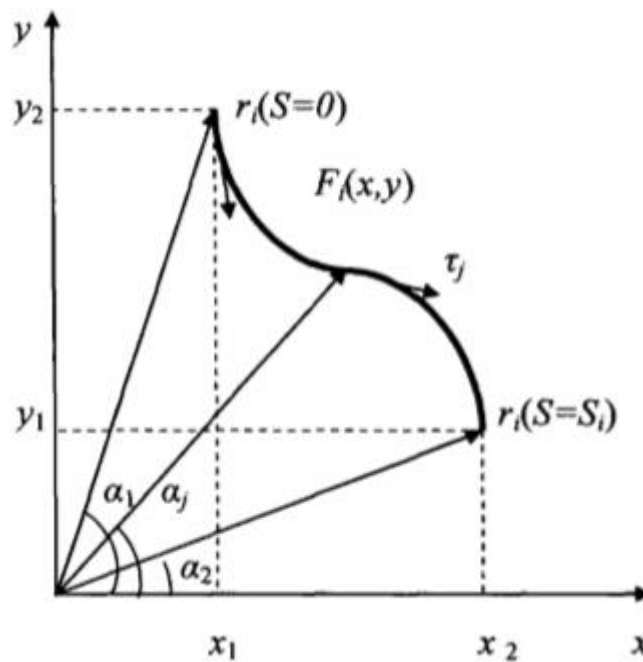


Рисунок 2.13 – Фрагмент кордону сегмента

Розглянемо довільну замкнуту криву, представлену на рисунку 2.14. Замкнута крива $F(x, y)=0$ представлена в файлі даних у вигляді послідовності координат пікселів (так як зображення бінарне, то значення самої функції у файлі даних можна не зберігати і описати функцію тільки своїми аргументами)

$$(x_k, y_k), k \in K \quad (2.31)$$

таким чином, що множина $\{(x_k, y_k)\}$ (є об'єднанням двох непересічних підмножин P_i та Q_i , тобто

$$\{(x_k, y_k)\}_i = P_i \cup Q_i, \quad (2.32)$$

де множина

$$P_i = \{(x_g, y_g)\}, \forall g \in G, \quad (2.33)$$

$$[G \in K_i, \inf(g) = \arg x_l(g) = l, \sup(g) = \arg x_N(g) = N],$$

в якому елементи упорядковані за зростанням x_g , а множина

$$Q_i = \{(x_l, y_l)\}, \forall l \in L, \quad (2.34)$$

$$[L \in K_i, \inf(l) = \arg x_N(l) = N, \sup(l) = \arg x_l(l) = 2N-2],$$

в якому елементи упорядковані за зменшенням x_l , при цьому в парі

$$\{(x_k, y_k)\}_i \in P_i \text{ та } \{(x_{N+(N-g)}, y_{N+(N-g)})\}_i \in Q_i, \quad y_{N+(N-g)} \geq y_g, \quad (2.35)$$

а в парі

$$\{(x_l, y_l)\}_i \in P_i \text{ та } \{(x_{N+(N-l)}, y_{N+(N-l)})\}_i \in Q_i, \quad x_l \geq y_{N+(N-l)}. \quad (2.36)$$

На рисунку 2.14 наведено приклад безлічі пікселів, яке утворено двома непересічними множинами P_i та Q_i які відповідають умовам (2.35) і (2.36) і створює замкнену криву – контур.

Так як не всяка крива є замкнутою і задовольняє умовам (2.35) і (2.36), то необхідний спосіб, що дозволяє нормалізувати дані так, щоб ці умови виконувалися для довільної кривої.

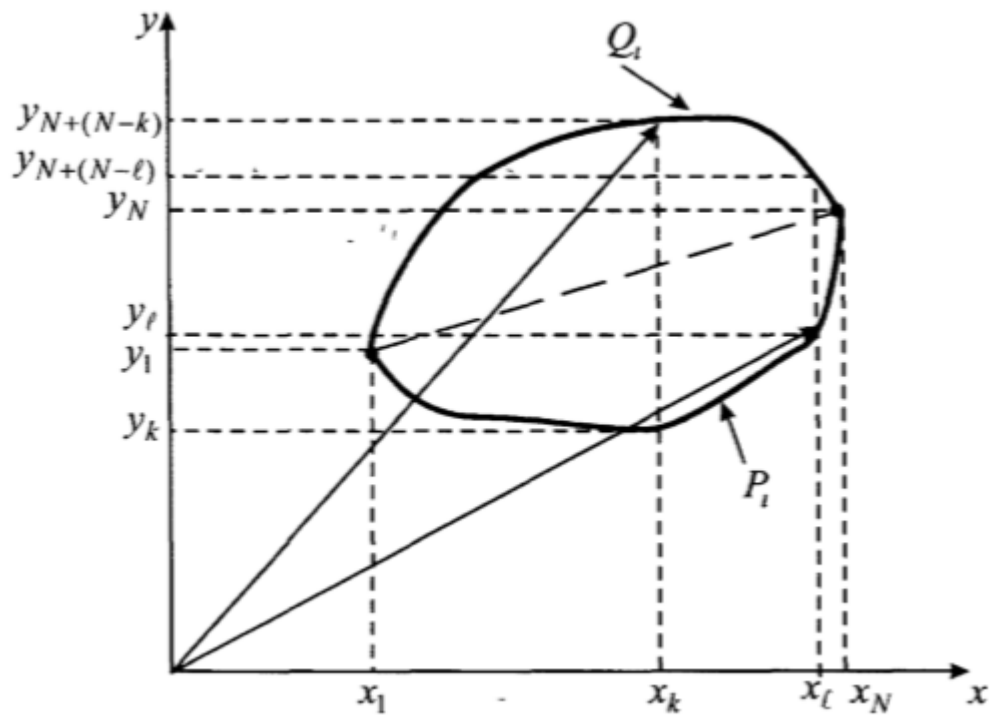


Рисунок 2.14 – Приклад розбиття, що задовольняє умовам (2.35) і (2.36), представляє контур у вигляді двох непересічних підмножин

Перший крок нормалізації – фіксування положення контуру в просторі декартових координат, тобто відбудова від перешкоди типу «зрушення». Один із способів вирішення цього завдання – зсув сегмента по осі x на величину x_1 і зрушення сегмента по осі y величину y_1 (рисунок 2.14).

Зрушення зображення сегмента здійснюється за допомогою наступного перетворення координат:

$$\begin{cases} \tilde{x}=x-x_1, \\ \tilde{y}=y-y_1. \end{cases} \quad (2.37)$$

Якщо крива $F(x, y)=0$, що описується можинами P_i та Q_i замкнута, вона є періодичною і може бути представлена у вигляді дискретного перетворення Фур'є.

Відліки спектру Фур'є отриманої замкнутої кривої використовуються для її параметричного представлення.

Як відліки розглядаємо комплексні відліки, отримані на дискретній сітці з рівномірним кроком по осі абсцис, визначувані як

$$\Delta = (x_N - x_1) / N, \quad (2.38)$$

де N – число відліків, що описуються межі сегменту, причому вважаємо, що число елементів у множині P_i і Q_i однаково, що відповідає умові

$$x_k = x_{N+(N-k)}, \quad \forall (x_k, y_k) \in P_i, \quad \forall (x_{N+(N-k)}, y_{N+(N-k)}) \in Q_i. \quad (2.39)$$

У окремому випадку в якості кроку дискретизації вибирається один піксель. При цьому необхідно враховувати масштаб початкового зображення і приводити крок дискретизації до масштабу (до реального зображення).

Об'єднання множин P_i і Q_i розглядаємо як масив впорядкованих комплексних чисел

$$\{\hat{z}_\zeta = x_\zeta + jy_\zeta\}, \quad \zeta = 1, 2, \dots, 2N-2, \quad (2.40)$$

для обробки якого можемо використати спектральні методи аналізу [29].

Криві лінії, що характеризують топологію сегментів зображення, не є замкнутими, тому не задовольняють умовам (2.38) і (2.40). Для замикання сегменту зображення розглянемо два підходи.

Перший підхід ґрунтований на представленні кривої в параметричній формі і дозволяє отримати замикаючий сегмент кривої за допомогою симетричного відображення початкових параметричних кривих. Його принцип ілюструє рисунок 2.14.

Метод такого перетворення реалізується досить просто. Необхідно масив даних, що містить координати x , збільшити в два рази, тобто до $2L_i$ і

записати в знову освічені осередки значення параметрів x в зворотному порядку:

$$x(L_I + l) = x(L_I - l), \quad (2.41)$$

де $l = L_I + 1, \dots, 2L_I$.

Для координати y використовується рівняння, аналогічне (2.41).

Однак такий підхід має істотний недолік, коли отримуємо характерну топологію у вигляді розгалуження. Для такої топології кривої скористатися рівняннями типу (2.40) не представляється можливим.

Тому розглянемо другий підхід, що припускає використання інтелектуальної обробки даних і отримання контурів мінімальної товщини (товщина в один піксель).

Цей підхід продемонстрований на рисунку 2.15, на якому показана деревовидна крива (рисунок 2.15, $x(l)$), перетворена в «контур мінімальної товщини».

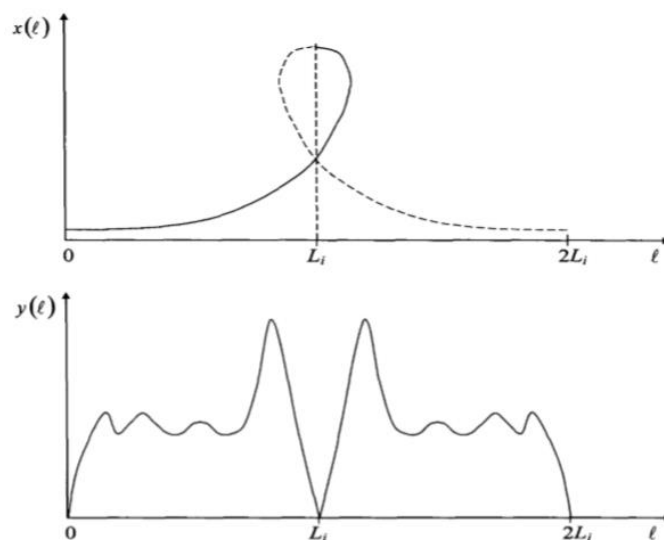


Рисунок 2.15 – Симетричні відображення відносно ординат, що дозволяють замкнути криву, задану в параметричній формі

2.4 Дослідження методів спектрального уявлення контурів мінімальної товщини

Розглянемо K -точечну дискретну межу «контура мінімальної товщини» на площині xoy , при цьому початкова точка (x_0, y_0) обрана довільно.

Починаючи з довільної точки (x_0, y_0) , здійснивши обхід контуру проти годинникової стрілки, отримаємо множину координат точок межі $\{(x_0, y_0), (x_1, y_1), \dots, (x_{k-1}, y_{k-1})\}$ – ці координати можна представити в параметричній формі

$$x_k = f_1(k), \quad (2.42)$$

$$y_k = f_2(k). \quad (2.43)$$

З використанням таких позначень кордон можна представити у вигляді послідовності координатних пар

$$s(k) = [f_1(k), f_2(k)], \quad (2.44)$$

де $k = \overline{0, K-1}$.

Кожну пару координат можна розглядати як комплексне число виду

$$z(k) = f_1(k) + jf_2(k), \quad (2.45)$$

для $k = \overline{0, K-1}$.

Таким чином, x і y розглядаються як дійсна та уявна частини для послідовності комплексних чисел.

Незважаючи на зміни спосіб інтерпретації цієї послідовності, сутність кордону залишилася колишньою, тобто вона описується тією ж послідовністю координат (2.42).

Таке уявлення має перевагу: він дозволяє звести двовимірну задачу спектрального аналізу до одновимірної [30].

Дискретне перетворення Фур'є кінцевої послідовності $z(k)$ задається виразом:

$$\hat{a}(u) = \frac{1}{2N-2} \sum_{k=0}^{N-3} \hat{z}(k) e^{\frac{j2\pi uk}{2N-2}}, \quad (2.46)$$

де $u = 0, 1, 2, \dots, 2N-3$.

Коефіцієнти $\hat{a}(u)$, що описують кордону, називають дескрипторами Фур'є. Зворотне перетворення Фур'є, застосоване до цих коефіцієнтів, дозволяє відновити межу z_k :

$$z(k) = \frac{1}{K} \sum_{u=0}^{K-1} a(u) e^{j2\pi uk/K}, \quad (2.47)$$

для $k = 0, 1, 2, \dots, K-1$.

Дескриптори Фур'є дозволяють синтезувати простір інформативних ознак в області просторових частот для ідентифікації форми контура, при цьому вони мають бути інваріантні відносно переміщення і повороту [31]. Інваріантність дескрипторів Фур'є відносно масштабу не обов'язкова, хоча зміна масштабу контура при рішенні завдань класифікації може бути як перешкодою, так і інформативним джерелом.

Дескриптори мають бути як можна менш чутливими до геометричних перетворень: зрушенню, зміні масштабу і повороту об'єктів. Для кожної інваріантності втрачається один ступінь свободи.

Для інваріантності відносно переміщення (дві інваріантності по x і y) не враховуємо перший дескриптор Фур'є $a(0)$ – втрачаємо два ступені свободи, дві координати).

Для інваріантності відносно масштабу встановлюємо абсолютну величину другого дескриптора Фур'є $a(l)$ рівній одиниці (один ступінь свободи), а для інваріантності відносно повороту співвідносимо усі фази з фазою $a(l)$ (ще один ступінь свободи).

При усіх трьох інваріантах втрачаються чотири ступені свободи.

У тих випадках, коли на результат обробки впливає порядок обходу точок межі (зустрічаються такі криві), встановлюється додаткова вимога незалежності дескрипторів від вибору початкової точки. Вираз $a_p(k)=a(k-k_0)$ означає перевизначення послідовності відповідно до рівняння

$$a_p(k)=x(k-k_0)+jy(k-k_0), \quad (2.48)$$

що відповідає зміні початкової точки послідовності з $k=0$ на $k=k_0$. Зміна початкової точки впливає на усі дескриптори по-різному, хоча відомим способом, оскільки $a(u)$ множиться на член, залежний від u .

Перевагою використання дескрипторів Фур'є для опису геометричних структур виділених сегментів є те, що викладені вище інваріанти містяться в двох перших дескрипторах Фур'є. Якщо нормуються інші дескриптори Фур'є за допомогою фази і абсолютної величини другого дескриптора Фур'є, то виходить повне інваріантне відносно переміщення, повороту і масштабу опис форми об'єктів.

Не враховуючи дескриптори Фур'є більш високого порядку, можемо поступово випускати дрібні деталі з опису форми керованим чином [32].

Інваріантність відносно здвигу. Положення об'єкту (контура) обмежується єдиним коефіцієнтом a_0 (нульовою гармонікою).

Усі інші коефіцієнти є інваріантними відносно переміщення. Інваріантність відносно перетворення «зрушення» можна досягти прирівнюванням до нуля нульового коефіцієнта.

Того ж ефекту можна досягти шляхом переміщення сегменту в центр декартових координат. Оскільки сегмент, в загальному випадку, не симетрична фігура, те розміщення його в центрі координат неоднозначно.

Досліджуваний сегмент вписується в прямокутник, і описаний прямокутник переміщається в початок координат. Використовуючи перетворення типу «зрушення», перераховуються усі координати сегменту, при обліку того, що зрушення здійснюється на відстань, що бере початок відліку від координат центра описаного прямокутника до центра координат.

Інваріантність відносно масштабу. Якщо контур масштабується коефіцієнтом a , те усі дескриптори Фур'є також масштабуються a . Якщо контур сегменту має ненульову площу і простежується проти годинникової стрілки, то перший коефіцієнт (перша гармоніка) завжди не дорівнює нулю.

Таким чином, масштабування дескрипторів Фур'є досягається шляхом їх нормування відносно абсолютної складової першого дескриптора Фур'є $|a_1|$. У таблиці 2.4, як приклад, приведені дані абсолютних складових перших п'яти дескрипторів Фур'є ($|a_i|$, де $i = -5, -3, \dots, 5$) для різних моделей сегментів центральних ліній кровоносних судин сітківки ока з різною довжиною (з різним числом відділів K).

Таблиця 2.4 – Дескрипторів Фур'є контурів різних розмірів

Число точок в контурі K	Дескриптори Фур'є (модуль комплексного ектора)										
	$ a_{-5} $	$ a_{-4} $	$ a_{-3} $	$ a_{-2} $	$ a_{-1} $	$ a_0 $	$ a_1 $	$ a_2 $	$ a_3 $	$ a_4 $	$ a_5 $
220	0,47	0,01	0,00	0,06	38,00	0,62	0,00	0,06	0,39	0,01	0,00
124	0,31	0,04	0,00	0,09	21,80	0,59	0,00	0,09	0,32	0,04	0,00
356	1,05	0,03	0,00	0,03	62,86	0,68	0,00	0,03	1,00	0,03	0,00
516	1,46	0,02	0,00	0,04	90,94	0,65	0,00	0,04	1,42	0,02	0,00
196	0,63	0,04	0,00	0,06	34,91	0,63	0,00	0,06	0,66	0,04	0,00

У таблиці 2.5 приведений результат нормування дескрипторів Фур'є відносно абсолютної складової першого дескриптора Фур'є.

Аналіз значень таблиці 2.5 показує, що дескриптори з молодшими індексами у сегментів однакової форми, але різного масштабу стали приблизно однакові.

Таблиця 2.5 – Нормованих дескрипторів Фур'є контурів різних розмірів

Число точок в контурі K	Дескриптори Фур'є (модуль комплексного ектора)										
	$ a_{-5} $	$ a_{-4} $	$ a_{-3} $	$ a_{-2} $	$ a_{-1} $	$ a_0 $	$ a_1 $	$ a_2 $	$ a_3 $	$ a_4 $	$ a_5 $
220	0,1	0,00	0,00	0,00	1,00	0,02	0,00	0,00	0,1	0,00	0,00
124	0,1	0,00	0,00	0,00	1,00	0,03	0,00	0,00	0,1	0,00	0,00
356	0,2	0,00	0,00	0,00	1,00	0,01	0,00	0,00	0,2	0,00	0,00
516	0,2	0,00	0,00	0,00	1,00	0,01	0,00	0,00	0,2	0,00	0,00
196	0,2	0,00	0,00	0,00	1,00	0,02	0,00	0,00	0,2	0,00	0,00

Різними залишилися дескриптори із старшими індексами, оскільки вони визначають той високочастотний сегмент спектру, в якому проявляються відмінності форми еритроцитів різних розмірів.

Забезпечення інваріантності відносно масштабу є актуальним при порівнянні і класифікації сегментів на різних фотографіях, зроблених з різними масштабами.

Інваріантність відносно повороту. Інваріантність відносно повороту задається абсолютною величиною дескрипторів Фур'є.

Це означає, що представлення дескрипторів у вигляді модулів комплексного вектору інваріантно відносно повороту. Отже, в завданнях класифікації необхідно використати тільки амплітудні значення спектру.

2.5 Вибір та оптимізація кількості дескрипторів Фур'є при аналізі кордонів сегментів складноструктурованих зображень

Для того, щоб система розпізнавання була адекватною, необхідно, щоб частотний діапазон, що відповідає дескриптору Фур'є з певним номером, не залежав від числа відліків у контурі, спектр якого обчислюється. Так як в загальному випадку в різних контурах, що визначають межі сегмента, міститься різне число відліків, то при формуванні інформативних ознак з спектральних відліків вони повинні відповідати одним і тим же частотними діапазонами. При однакової частоті дискретизації і різних кількостях відліків в контурах ця вимога порушується. Додаткові відліки можна отримати в результаті інтерполювання замкнутої кривої, що відповідає контуру. Але в нашому випадку такий шлях неприйнятний, так як контур дискретизований з граничною частотою дискретизації, тобто один піксель.

У зв'язку з вищевикладеним, був запропонований наступний метод формування простору інформативних ознак.

На першому етапі визначаємо загальне число відліків в контурах, яке повинно бути однаковим для всіх контурів контрольних і навчальних вибірок. Це число визначається за результатами статистичних досліджень.

Для того щоб мати можливість порівнювати дискретні відліки частот, що відповідають різним контурам, необхідно, щоб контури мали одну і ту ж кількість відліків (априорно вважається, що частота дискретизації у них однакова і дорівнює одиниці, тобто один піксель).

Щоб вирівняти число відліків у вибірці аналізованих контурів, необхідно довести число відліків в кожному контурі до максимального у вибірці.

Це можна здійснити шляхом багаторазового обходу контуру, тобто його періодичного продовження до тих пір, поки число відліків не досягне максимуму за вибіркою або максимально наблизиться до нього.

Однак, такий спосіб вирівнювання відліків у вибірці вимагає, щоб у вибірці всі контури мали кратне число відліків. Практично ця вимога, що не реалізується.

Можна задати число відліків значно більше, ніж максимальне у вибірці, наприклад, у десять разів, і тоді енергія спектру, пов'язана з помилковими гармоніками, значно зменшиться. Але таке технічне рішення призведе до затягування обчислювального процесу, що небажано при роботі в інтерактивному режимі, що передбачає обчислення спектру контурів при різних значеннях параметрів.

Відомо, що інтерполяції у просторі сигналів відповідає збільшення смуги частот, що займається сигналом, в спектральній області [33].

Отже, якщо доповнити високочастотну частину спектральної смуги нулями і тим самим довести число відліків в спектрі кожного контуру межі сегмента до максимального, то в просторі сигналів з'являться віртуальні відліки між реальними відліками.

Таким чином, маємо наступний ланцюжок перетворень: $K_{max} \rightarrow$ спектр $\rightarrow$ доповнення нулями спектральних відліків в області високих частот до K_{max} .

Проте характерною особливістю дескрипторів Фур'є є те, що їх амплітуда пов'язана з частотою.

Тому будь-які частотні морфізми в реальному сигналі призводять до амплітудним змін спектральних складових. Критерієм адекватності будь-яких морфізмів у частотній області служить зворотне перетворення Фур'є і відповідні відмінності між прямим і зворотним перетворенням Фур'є.

Приклади, пов'язані з відповідністю між перетворенням частоти і амплітуди сигналу, наведені в [33].

Доповнення нулями спектра відповідає інтерполяції у просторі сигналів, тобто появи додаткових відліків між вихідними відліками.

Так як контур дискретизований з граничною частотою дискретизації, то проміжні відліки потрапляють у ті самі координати, що і вихідні.

Якщо спектр визначається на контурі, то координати відліку визначають величину відліку. Тоді додаткові відліки мають значення, що дорівнюють значенням у суміжних вузлах інтерполяції. Це веде до порушення вимог рівності Парсеваля.

Щоб зберегти енергетичний еквівалент між простором сигналів і простором частот, необхідно зменшити амплітуду дескрипторів пропорційно числу проміжний час, що з'явилися в просторі сигналів в результаті доповнення спектру нулями.

Таким чином, на другому етапі необхідно все спектральні складові в спектрі i -го контуру помножити на величину K_{max}/K_i , де $K_{max}=3000$, K_i – кількість відліків у i -му контурі.

Третій етап – оптимізація кількості використовуваних дескрипторів. Для оптимального вибору числа аналізованих дескрипторів (під оптимальністю тут розуміється мінімізація їх числа) необхідно здійснити зворотні перетворення Фур'є модифікованого спектру контуру і порівняти його з вихідним контуром.

Враховуючи, що перетворення Фур'є оборотно, і використовуючи рівняння (2.48), за дескриптором Фур'є можемо відновити межі контуру досліджуваного сегмента. У будь-якому випадку, для відновлення контуру використовуються стільки дескрипторів, скільки було отримано відліків на контурі. Якщо частина дескрипторів прирівняти до нуля, то, при використанні нейромережевої класифікаційної моделі, вузли вхідного шару нейронної мережі, відповідні цим дескрипторам, можна виключити, що призводить до значного спрощення моделі прийняття рішення.

Щоб оцінити інформаційні втрати при прирівнювання до нуля частини дескрипторів, необхідно порівняти за певним критерієм вихідний контур і контур, відновлений за обмеженого набору дескрипторів. З цією метою представимо контур двома параметричними кривими (2.47) і (2.48).

Після переходу від K дескрипторів до M дескрипторів ($M < K$), можемо отримати параметричні криві, що відображають геометрію межі сегмента у вигляді

$$\tilde{x}_k = \tilde{f}_1(k), \quad (2.49)$$

і

$$\tilde{y}_k = \tilde{f}_2(k). \quad (2.50)$$

Приклади вихідних та відновлених з M дескрипторів параметричних кривих, (2.30) і (2.47), (2.48) представлені на рисунку 2.16.

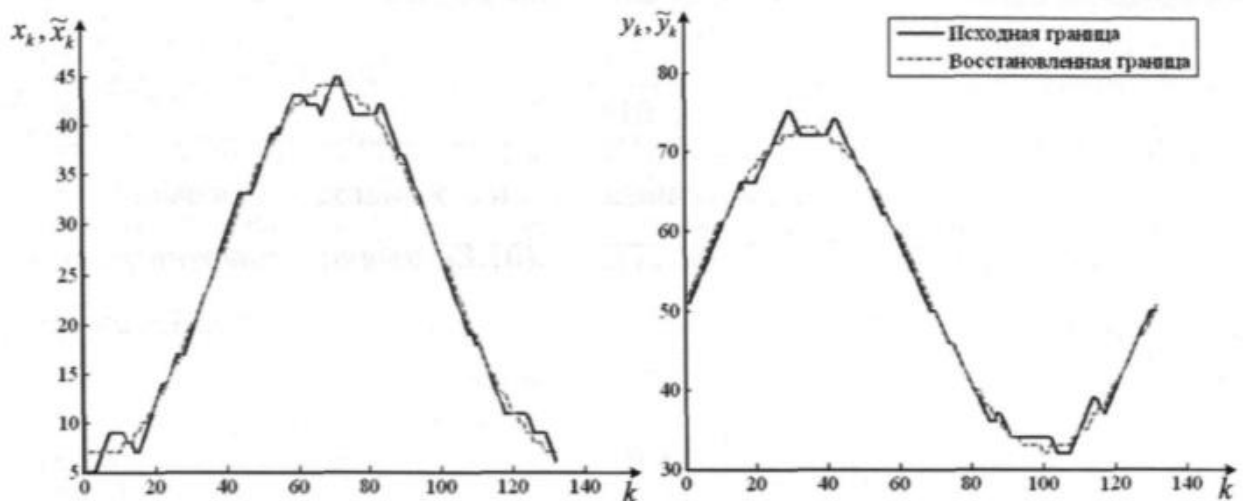


Рисунок 2.16 – Параметричні криві вихідної і відновленої межі контуру

Так як параметр кривих, приклади яких наведені на рисунку 2.16, один і той же, то площа S контуру мінімальної товщини визначається як

$$S = \frac{1}{K} \sum_{k=0}^{K-1} |y(k(\xi, \zeta)) - y(k^*(\xi^*, \zeta^*))| \rightarrow \min, \quad (2.51)$$

де K – число відліків в контурах вибірки;

ξ^* та ζ^* – значення параметра k^* (відновленого контуру), знаходиться на мінімальній відстані від k (вихідного контуру), причому для растрових зображень $k-k^*$.

Для «контурів мінімальної товщини»

$$k^* = \text{int}(K/2) + k. \quad (2.52)$$

Точне визначення координат k^* здійснюється за допомогою установки маски розміром 3×3 (або 4×4) в піксель з номером до таким чином, щоб пікселі контуру увійшли в маску.

Піксель k^* належить масці і є найближчим до нормалі, відновленої з пікселя до граничної кривої, що ілюструє рисунок 2.17.

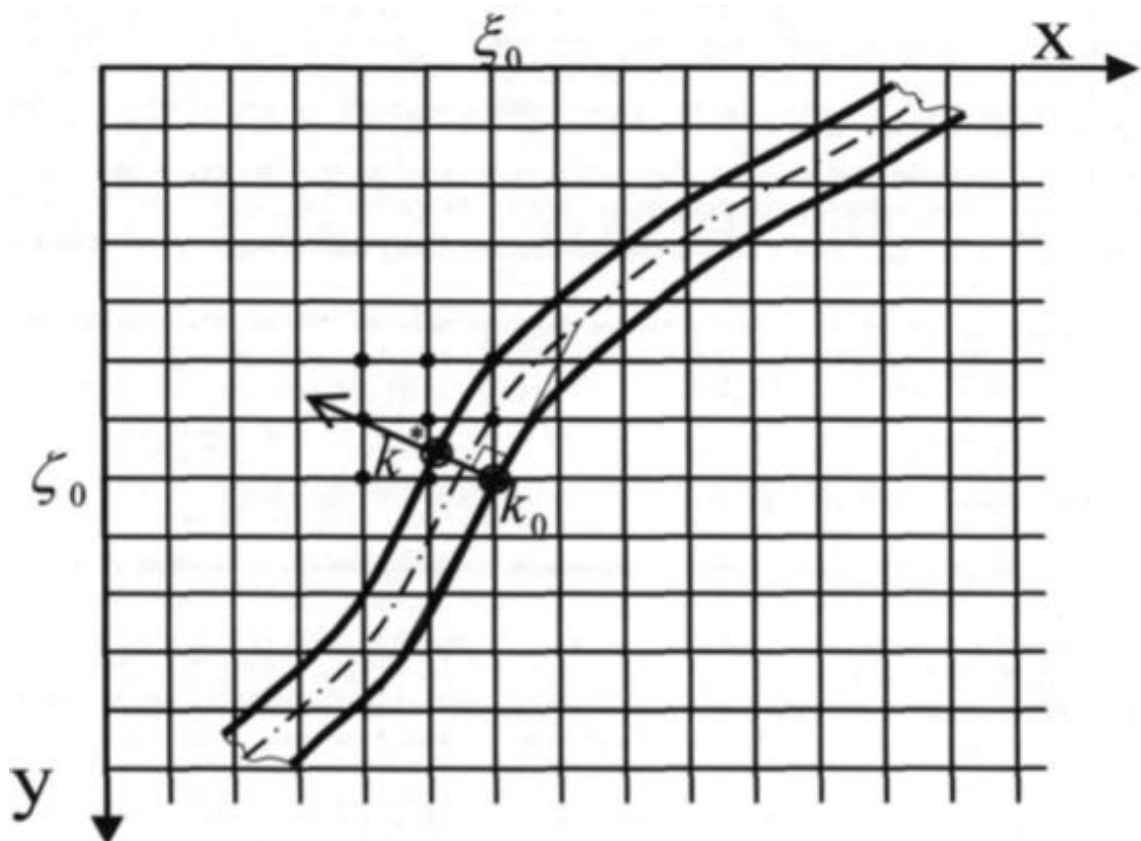


Рисунок 2.17 – Вибір координат відліків k^*
для контурів мінімальної товщини

При відсутності помилки площа $S \approx 1$, відповідно (2.52). Виконання цього наближеної рівності є індикатором того, що контур належить до класу контурів мінімальної товщини. Як критерій помилки відновлення, що оптимізується, використовуємо параметр A , що визначається як

$$A = S - I \rightarrow 0. \quad (2.53)$$

На рисунку 2.18 представлені графіки залежності інформаційних втрат від числа дескрипторів, що використовуються при відновленні межі контуру сегмента, для двох контурів.

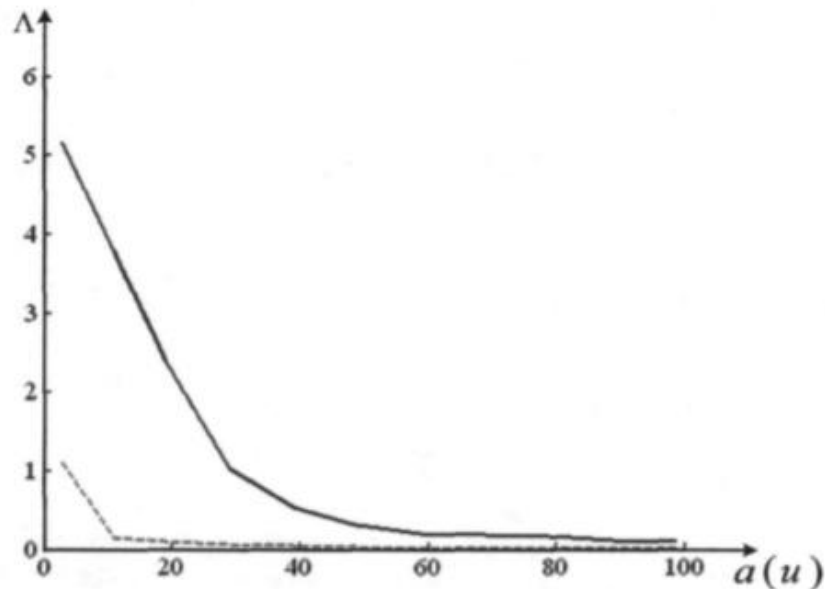


Рисунок 2.18 – Графік залежності інформаційних втрат від числа дескрипторів для двох контурів

Для оптимізації числа дескрипторів Кеттела [34] запропонував знайти таке місце на графіках, де спадання критерію A зліва направо максимально сповільнюється. Передбачається, що праворуч від цієї точки знаходиться тільки «інформативна осип» – «осип» є геологічним терміном, що позначає уламки гірських порід, які скупчуються в нижній частині скелястого схилу.

3 МЕРЕЖЕВІ ІНТЕЛЕКТУАЛЬНІ ТЕХНОЛОГІЇ ДЛЯ КЛАСИФІКАЦІЇ СЕГМЕНТІВ НА СКЛАДНОСТРУКТУРОВАНИХ ЗОБРАЖЕННЯХ

3.1 Універсальна мережева структура для задачі класифікації

В даний час існують численні варіанти нейронних мереж, що володіють властивостями універсальних аппроксиматорів функцій багатьох змінних. Нейронні мережі побудовані таким чином, що можуть обчислювати лінійні функції, нелінійні функції однієї змінної, а також всілякі суперпозиції функцій, одержувані при каскадному з'єднанні мереж.

Робота такого аппроксиматора ґрунтується на теоремі Арнольда-Колмогорова, згідно якої будь-яка безперервна функція f змінних, задана на одиничному кубі n -вимірного простору $I^n = I \times I \times \dots \times I$ може бути представлена у вигляді

$$f(x_1, x_2, \dots, x_n) = \sum_{q=1}^{2^{n+1}} h_q \left[\sum_{p=1}^n \phi_q^p(x_p) \right], \quad (3.1)$$

де функції h_q неперервні функції однієї змінної, а функції $\phi_q^p(x_p)$ – безперервні незменшувані на $I=[0,1]$ стандартні функції.

Для реалізації (3.1) до складу нейронної мережі повинні входити нелінійний перетворювач сигналів, нелінійний адаптивний суматор і блок галуження, що здійснює розсилку одного і того ж сигналу по декількох адресах. Нейронна мережа, побудована згідно з рівнянням (3.1), є неоднорідною нейронною мережею.

На рисунку 3.1 представлена модель універсальної структури нейронної мережі, побудована на основі рівняння (3.1), в просторі двох змінних.

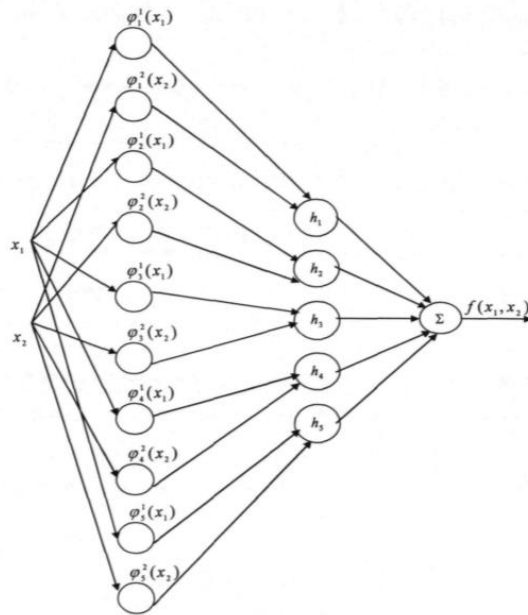


Рисунок 3.1 – Модель структури універсальної неоднорідної нейронної мережі для двовимірного простору інформативних ознак

Універсальний апроксиматор, представлена формулою (3.1), можна поліпшити. Зокрема [34], функції h_q можна замінити на одну універсальну безперервну функцію h , яку можна представити у вигляді ряду, що абсолютно сходиться, Фур'є, а функції $\phi_q^p(x_p)$ – можна «розщепнути» на похідні $l_p \phi_q$, ($l_p, p=1, 2, \dots, n$) – раціональні незалежні числа, а $\phi_q, (q=1, 2, \dots, 2n+1)$ – безперервні незменшувані функції на $I=[0, 1]$.

Тоді, подання будь-якої неперервної функції $f \in C[I^n]$ можна записати у вигляді:

$$f(x_1, x_2, \dots, x_n) = \sum_{q=1}^{2n+1} h[\sum_{p=1}^n l_p \cdot \phi_q(x_p)]. \quad (3.2)$$

На рисунку 3.2 показана мережева структура, відповідна формулі (3.2). Вона являє багат шарову нейронну мережу, яку складно реалізувати на практиці. Основна проблема полягає в тому, що функції ϕ_q в формулі (3.2) неперервні проте не є гладкими.

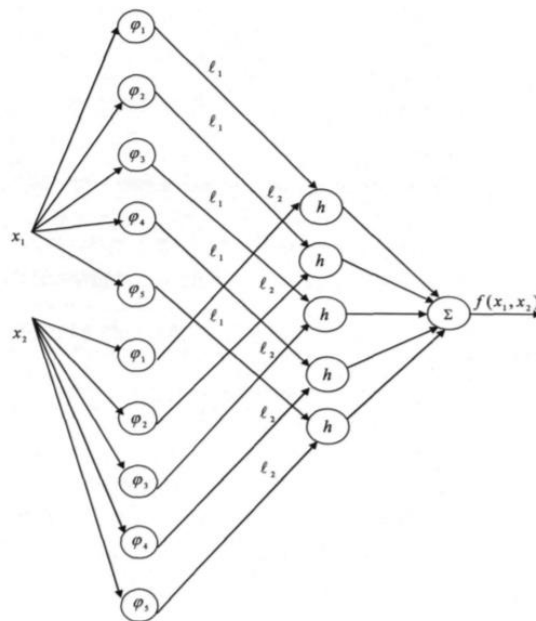


Рисунок 3.2 – Модель універсальної неоднорідною нейронної мережі для двовимірного простору інформативних ознак з модифікованою структурою

Оскільки довжина машинного коду функції обернено пропорційна до її гладкості, отримаємо дуже погану точність реалізації колмогоровського представлення. Крім того, мережеве представлення повинне мати фіксовані елементи і змінювані параметри. Мережа, представлена на рисунку 3.2, не є мережею такого типу: форма h залежить від функції f , яку необхідно представити (функції ϕ_q не залежать від f) [35].

Основною проблемою, пов'язаною з реалізацією мережевих структур, показаних на рисунку 3.1 і рисунку 3.2, є те, що вони неоднорідні, тобто не мають фіксованих елементів і параметрів, що настраюються: функції h у вузлах мереж залежать від функції, що апроксимується f .

Щоб спростити процес апроксимації, замінимо досить складну гіперповерхню, відповідну рівнянь (3.1) і (3.2), гіперплощиною, описується виразом

$$f(x_1, x_2, \dots, x_n) = a_1 x_1 + a_2 x_2 + \dots + a_n x_n - b, \quad (3.3)$$

де a_i та b – константи ($i = \overline{1, n}$).

На рисунку 3.3 представлена мережева структура, що реалізує описаний спосіб апроксимації, в якій вихідна система (3.7) розщеплена на λ непаралельних гіперплоскостей.

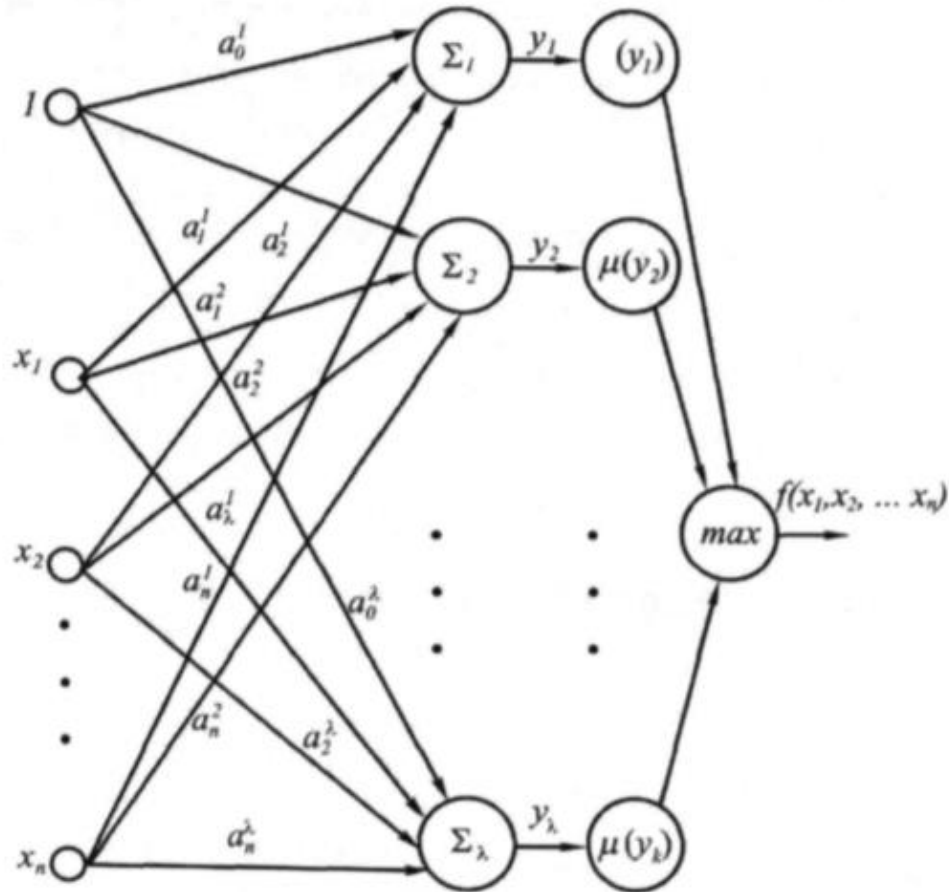


Рисунок 3.3 – Мережева структура універсального класифікатора

Мережа складається з чотирьох шарів і є однорідною, так як у вузлах мережі здійснюються однакові математичні операції. Перший (вхідний) шар визначає вектор стану системи класифікації. Другий шар синтезує апроксимуючі функції згідно рівнянню (3.3).

Число λ таких апроксимуючих функцій визначається структурою даних – навчальною вибіркою і апріорно невідомо.

Вибір величини ε визначається відстанню між класами. Якщо прийняти, що відстань між діагностуються класом ϑ і альтернативним класом $\overline{\vartheta}$ дорівнює

одиниці і всі об'єкти цих класів розміщені всередині гіперсфер діаметром одиниця, то величину ϵ визначаємо з рівняння

$$\epsilon = \frac{1}{\lambda + 1}. \quad (3.9)$$

З формули (3.9) випливає, що вибір значення ϵ представляє компроміс між збільшенням числа нейронів у другому шарі і збільшенням точності апроксимації.

Так як мережа, представлена на рисунку 3.3 налаштована на один клас – ϑ , то при поданні на її входи компонентів вектора X_j , відповідних цьому класу, на виходах нейронів другого шару з'являться числа, близькі до нуля. При цьому, найбільш близьким до нуля буде вихід того нейрона, відстань між гіперплощиною якого і точкою X_j мінімально.

Після цього досить вибрати мінімальне значення з усіх компонентів вектора стану другого шару $(y_1, y_2, \dots, y_r, \dots, y_\lambda)$, яке буде характеризувати ступінь близькості вектора X до класу ϑ .

Так як вихідні сигнали другого шару можуть наближатися до нуля з двох сторін, то над ними повинні бути здійснені симетричні нелінійні перетворення. Ці нелінійні перетворення дозволяють здійснити перехід від чітких чисел, відповідних виходам другого шару, до нечітким, які характеризують ступінь приналежності відповідного виходу другого шару до потрібного класу. Так як ці виходи однозначно пов'язані (через відповідні гіперплощини) з вхідним вектором, нелінійні перетворення назовемо функціями приналежності до шуканого класу базової змінної, що відповідає виходу другого шару.

В частому випадку, ці нелінійні перетворення не залежать від виходу нейрона другого шару і їх приклади показані у вигляді графіків на рисунку 3.4.

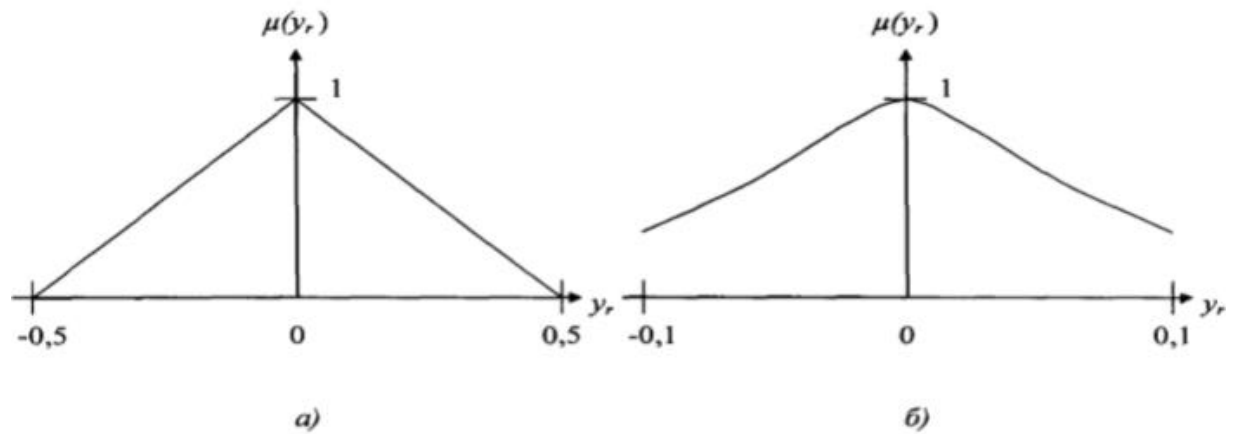


Рисунок 3.4 – Нелінійні перетворення виходу r нейронів другого шару до класу ϑ :

- а) до перетворення;
- б) після перетворення

У загальному випадку нелінійні перетворення отримують в процесі навчання для кожного нейрона або задаються емпірично.

При використанні нелінійних перетворень виду, показаного на рис. 3.4, в четвертому шарі мережі рисунок 3.3 здійснюється операція «нечітке або»

$$\max(\mu_1(y_1), \mu_2(y_2), \dots, \mu_r(y_r), \dots, \mu_\lambda(y_\lambda)) = f(x_1, x_2, \dots, x_n). \quad (3.10)$$

У практичних додатках використовуються два аппроксиматора, що відповідають двом розділяються класам. Це означає, що універсальна нейронна мережа для класифікації даних має блочну структуру, причому число блоків має відповідати числу поділюваних класів.

У відповідності зі структурою, представленою на рисунку 3.3 і рівняннями (3.7), (3.10), а також позначивши функцію «нечітке або»

$$\max(\mu(y_1), \mu(y_2)) = \sum_{i=1}^2 [\Theta] \mu_i(x), \quad (3.11)$$

а виходи нейронів другого шару

$$y_r = \sum_{i=0}^n a_i x_i \quad (3.12)$$

модель, що апроксимує для мережевої структури рисунка 3.3 висловимо наступним рівнянням

$$f(x_1, x_2, \dots, x_n) = \sum_{i=1}^k [\Theta] \mu_i [\sum_{i=0}^n a_i x_i] \quad (3.13)$$

або в більш загальному випадку

$$f(x_1, x_2, \dots, x_n) = \sum_{i=1}^k [\Theta] \mu_i [\sum_{i=0}^n \phi_i(x_i)]. \quad (3.14)$$

Рівняння (3.14) має запис, аналогічну рівнянням (3.1) і (3.2), відповідним теоремі Арнольда-Колмогорова, але позбавлене основних їх недоліків: всі функції, що входять в нього, не залежить від функції, що апроксимується.

3.2 Нейромережеві структури для сегментації зображень за параметрами пікселів та їх околу

Одним з атрибутів пікселя є його яскравість або RGB-код у разі кольорового зображення. Тому найбільш простим і поширеним методом сегментації є порогова або двухпорогова обробка, призводить до бінаризації зображення [37].

Фрагменти на складноструктурованому зображенні знаходяться в зонах різної освітленості, що потребує відбудови від фонові складові, тому для сегментації таких зображень використовуються різні градієнтні методи.

Вихідні атрибути пікселя є релевантними показниками при сегментації зображень, вони успішно використовуються в методах «розщеплення і злиття» або в методах нарощування областей» [38].

З метою зниження помилок при сегментації зображень доцільно використовувати адаптивні методи, в залежності від типу виділяється сегмента.

Так як однією з особливостей складноструктурованих зображень є різноманіття властивостей виділених сегментів і зашумленість їх меж, то при їх сегментації доцільно застосовувати гібридні методи або одночасно використовувати методи, засновані на різних парадигмах.

Для реалізації такого підходу використовуємо класифікують нейронні мережі, на входи яких подаються атрибути околиці аналізованого пікселя (околиця визначається апріорно заданої маскою, а в окремому випадку це один піксель), а на виході отримуємо сигнал, що визначає приналежність пікселя до заданого класу.

Залежно від конкретного об'єкта дослідження та обраного методу виділення сегмента можна запропонувати множина реалізацій нейронних мереж, що відповідають запропонованим походу.

Розглянемо базовий варіант нейромережевої структури, орієнтований на один піксел з урахуванням того, що ускладнення околиці і методу формування області сегмента не призводить до якісної зміни нейромережевої структури, а призводить лише до збільшення нейронів вхідного шару і ускладнення блокової структури.

Для поділу пікселів зображення очного дна на два класу «сегмент» і «сегмент» пропонується використовувати мережеву структуру, представлену на рисунку 3.5.

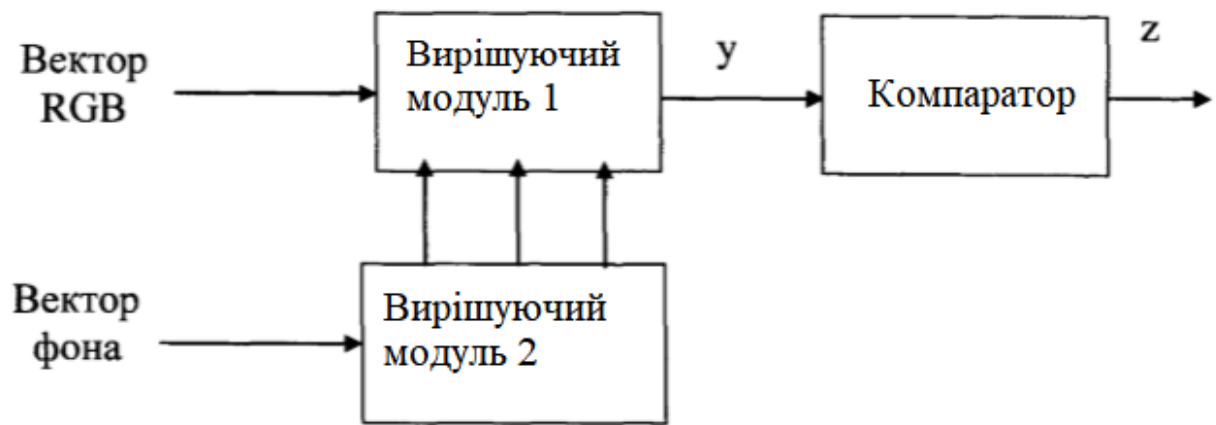


Рисунок 3.5 – Структурна схема мережевої моделі
для класифікації пікселів на зображенні

На вхід мережі подаються два вектора інформативних ознак. Перший трикомпонентний вектор інформативних ознак подається на перший вирішальний модуль і являє собою RGB-коди пікселів, які необхідно класифікувати.

На другий вирішальний модуль подається трикомпонентний вектор фону, значення компонент якого визначаються усередненими характеристиками RGB-кодів в околиці аналізованого пікселя, форма і розміри якої залежать від виду сегмента (задачі).

Другий вирішальний модуль здійснює управління структурою першого вирішального модуля. Це здійснюється наступним чином.

Перший вирішальний модуль налаштовується на клас «сегмент». Його структура відповідає універсальній мережній структурі рівняння (3.14). Згідно з цим апроксимующому рівняння повинно бути здійснено нелінійне відображення вхідного вектора у простір ознак тієї ж чи більш високої розмірності, доцільність якого доводить теореми Ковера [38].

Специфічність аналізованого класу зображень полягає в тому, що визуализируемая сцена на складноструктурованому зображенні розташовується на сферичній поверхні.

Тому RGB-коди не дозволяють відокремити сегмент від фону або інших сегментів зображення. Отже, необхідно ввести додаткові інформативні ознаки, що відображають рівень освітленості в околиці досліджуваного пікселя, або змінювати вид використовуваного нелінійного відображення в залежності від рівня освітленості, маючи при цьому певною кількістю рівнів освітленості і для кожного рівня побудувати відповідне нелінійне відображення.

Модель структури першого вирішального модуля представлена на рисунку 3.6.

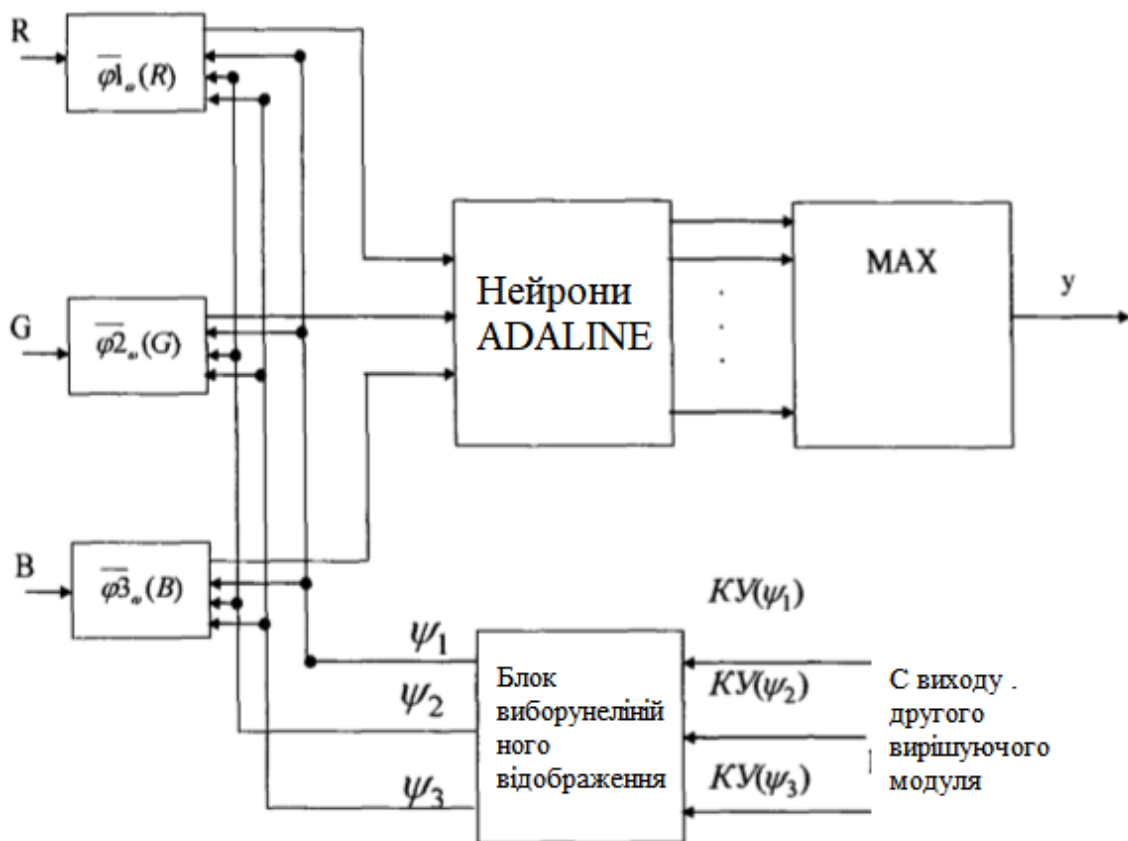


Рисунок 3.6 – Модель структури першого вирішального модуля

Як приклад на ній показаний вирішальний модуль, призначений для виділення сегментів на зображенні при поділі рівня яскравості (фону) в околиці досліджуваного пікселя на три класи.

Для кожного класу фону ψ_i формуються три нелінійних перетворення

$$[\phi_{\omega}^{\psi_k}(R), \phi_{\omega}^{\psi_k}(G), \phi_{\omega}^{\psi_k}(B)], k=1,2,3 \quad (3.15)$$

У моделі структури, представленій на рисунку 3.6, використовуються три класи фону ψ_1, ψ_2, ψ_3 коефіцієнтами впевненості для аналізованого пікселя $KU(\psi_1), KU(\psi_2), KU(\psi_3)$, які обчислюються у другому вирішальному модулі (рисунк 3.5).

У такому випадку для кожного інформативного ознаки (або вектора інформативних ознак, в загальному випадку) необхідно побудувати три нелінійних відображення, що відповідають трьом класам фону

Для формування апроксимуючих площин використовувалися нейрони adaline з функцією активації у вигляді, представленому на рисунку 3.4 а) чи 3.4 б).

$$\bar{\varphi}_{\omega}(R) = [\phi_{\omega}^{\psi_1}(R), \phi_{\omega}^{\psi_2}(R), \phi_{\omega}^{\psi_3}(R)]; \quad (3.16)$$

$$\bar{\varphi}_{\omega}(G) = [\phi_{\omega}^{\psi_1}(G), \phi_{\omega}^{\psi_2}(G), \phi_{\omega}^{\psi_3}(G)]; \quad (3.17)$$

$$\bar{\varphi}_{\omega}(B) = [\phi_{\omega}^{\psi_1}(B), \phi_{\omega}^{\psi_2}(B), \phi_{\omega}^{\psi_3}(B)]. \quad (3.18)$$

Блок MAX присвоює вихідної змінної z величину, що відповідає максимальному значенню виходів нейронів adaline.

Блок вибору нелінійного відображення підключає до вхідного вектора то нелінійне відображення, яке відповідає класифицируемому фону.

Для отримання нелінійних відображень (3.16–3.18) використовувалися гістограми розподілу інтенсивності RGB-кодів в пікселях зображень і експертні оцінки.

Другий вирішальний модуль будуємо за схемою багат шарових нейронних мереж прямого поширення [39].

Для класифікації фону розроблені дві моделі побудови другого вирішального модуля: неблокова структура (НБС) і блокова структура (БС).

Для вирішальної системи в неблочной структурі побудована одна нейронна мережа для класифікації всіх класів. Кількість нейронів вихідного шару нейронної мережі НБС дорівнює кількості класів.

Для реалізації мережевої моделі класифікації пікселів на складноструктурованому зображенні з метою визначення їх належності до заданого сегменту, структурна схема якої представлена на рисунку 3.5, розроблено відповідне методичне та програмне забезпечення.

4 ДОСЛІДЖЕННЯ ГІБРИДНИХ МЕТОДІВ ДЛЯ КЛАСИФІКАЦІЇ СКЛАДНОСТРУКТУРОВАНИХ ЗОБРАЖЕНЬ

4.1 Вибір інструментального засобу для вирішення поставлених задач

Для реалізації наступних алгоритмів було використано Python 3.6 та набір бібліотек cv2, numpy, skimage, scipy, math та datetime.

Python дозволяє швидко та ефективно реалізувати методи, що були розглянуті в данній роботі. На відміну від таких мов як C#, Java, C та C++, Python дозволяє швидко налаштувати оточення для розробки та реалізувати необхідні методи з мінімальною витратою часу та максимальною наочністю.

До того ж з кожним роком Python доповнюється новими бібліотеками з методами, що стануть у пригоді вченим та ентузіастам, що займаються обробкою зображень, AI та ML (англ. Artificial intelligence – штучний інтелект, Machine Learning – машинне навчання).

Для останніх двох пунктів підійде також мова R, проте на мою думку Python є більш простим та зручним для розуміння та використання.

Щоб мати можливість маніпулювати зображенням як параметром було використано бібліотеку cv2 – OpenCV для Python. Вона містить у собі достатній функціонал для роботи з візуальними об'єктами.

Для перетворення зображення у масив значень були використані функції бібліотеки numpy.

Щоб отримати доступ до необхідних методів обробки даних, отриманих за допомогою вищезазначених бібліотек у проект була імпортована бібліотека scipy.

Інші бібліотеки були допоміжними у конвертації та модифікації даних, їх коректної ілюстрації.

4.2 Вихідні дані для дослідження

Як вихідні дані для дослідження обрано зображення, що подане на рисунку 4.1.



Рисунок 4.1 – Вхідне зображення для тестування методів

Рисунок 4.1 має досить складну та навантажену структуру для виділення контурів та сегментів, тому його можна назвати складноструктурованим, а це оптимальний варіант для тестування.

Далі це зображення буде оброблено за допомогою ЕМ-метода, метода Канні та операторів Робертса і Собела. Порівнявши результати, отримаємо змогу оцінити роботу кожного з них і обрати найоптимальніший, або ж просто зробити деякі висновки щодо їх роботи.

4.3 Тестування гібридних методів для класифікації складноструктурованих зображень

Перший метод для аналізу зображення буде ЕМ-методу. ЕМ-метод використовується в математичній статистиці для знаходження оцінок максимальної правдоподібності параметрів імовірнісних моделей, в разі, коли модель залежить від деяких прихованих змінних. Кожна ітерація алгоритму складається з двох кроків. На Е-кроці (expectation) обчислюється очікуване значення функції правдоподібності, при цьому приховані змінні розглядаються як спостережувані. На М-кроці (maximization) обчислюється оцінка максимального правдоподібності, таким чином збільшується очікуване правдоподібність, що обчислюється на Е-кроці. Потім це значення використовується для Е-кроку на наступній ітерації.

Рисунок 4.2 показує передобробку зображення перед обробкою.



Рисунок 4.2 – Вихідне зображення (після обробки програмою).

Отриманий результат на рисунку 4.3. Аналізуючи зображення, можемо чітко бачити виділені сегменти, що утворились після обробки ЕМ-методом.

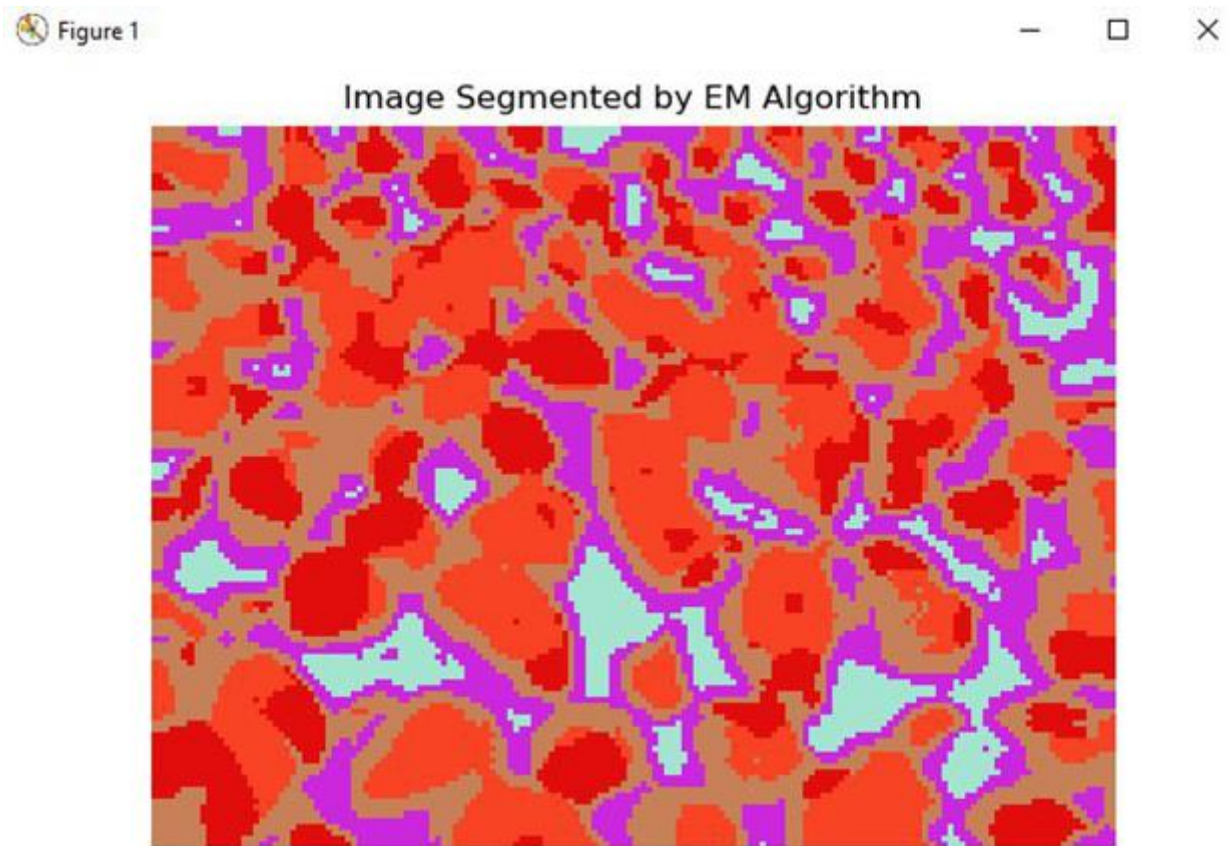


Рисунок 4.3 – Зображення після обробки (червоним виділені сегменти)

Проведемо аналогічний аналіз за допомогою метода Канні. Метод Канні відрізняється від інших градієнтних методів тим, що при визначенні меж використовує два порога (для слабких і сильних кордонів).

Слабкі кордони відзначаються в результуючому зображенні тільки тоді, коли вони з'єднані з сильними. Для зашумлених зображень даний метод забезпечує найкраще виявлення кордонів в порівнянні з іншими методами, але вимагає істотно більшого часу [40].

Згідно даного методу спочатку виконується згладжування зображення за допомогою гауссовського фільтра. Потім визначається величина стрибка (модуль градієнта) інтенсивності і його напрямок в кожній точці згладженого зображення.

Зображення диференціюють за двома ортогональними напрямками, потім обчислюють величину модуля градієнта як квадратний корінь з суми квадратів похідних. Майбутні точки контуру визначаються в результаті придушення точок, які не є локальними максимумами величини градієнта (non-maximal suppression).

У цьому випадку величина стрибка інтенсивності в кожній точці контуру покладається рівною нулю, якщо вона не перевищує величину стрибка інтенсивності в двох сусідніх точках по напрямку градієнта. Потім виконується порівняння величини інтенсивності в отриманих точках контуру з двома порогами методом гистерезиса.

Відповідно до цього методу точки, величина інтенсивності яскравості в яких менше нижнього порога, не включаються до контурний препарат. Точки, величина інтенсивності в яких вище нижнього порога, приєднуються до сусідніх точок, величина інтенсивності в яких перевищує верхній поріг.

Результати обробки зображення методом Канні проілюстровано на рисунку 4.4

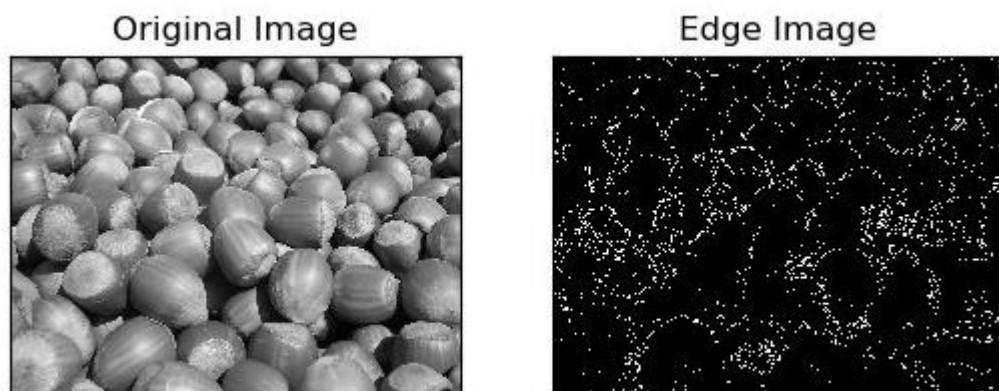


Рисунок 4.4 – Результат роботи методу Канні

Як можна бачити були виділені основні кордони на зображенні. Також можна задавати щільність точок, які виділяють кордони, проте якщо цей показник буде занадто високим (наприклад більше 100) отримаємо повністю біле зображення.

Проаналізуємо зображення за допомогою операторів Робертса і Собеля.

Оператор Робертса є найбільш простим і швидким методом виділення контурів. Його особливістю є інваріантність до обертання.

Оператор Собеля обчислює наближене значення градієнта яскравості зображення. Результатом застосування оператора Собеля в кожній точці зображення є або вектор градієнта яскравості в цій точці, або його норма.

На рисунку 4.5 зображені результати обробки зображення обома методами.

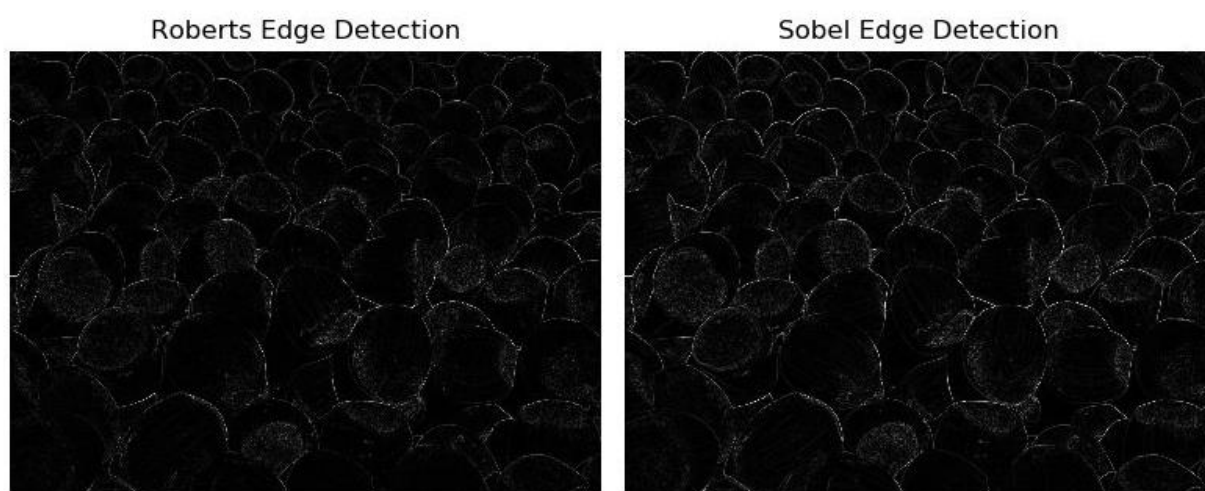


Рисунок 4.5 – Результат роботи операторів Робертса і Собеля

Як можемо бачити, робота обох операторів майже однакова, кордони визначені достатньо чітко.

4.4 Результати оцінювання ефективності гібридних методів для класифікації складноструктурованих зображень

Провівши аналіз результатів, можна із впевненістю сказати, що усі обрані методи впорались із поставленим завданням достатньо добре, тож визначити найкращий доволі складно, можливо навіть некоректно.

Для окремих завдань необхідно брати конкретні методи. Якщо це визначення кордонів то я обирав би оператор Собеля або оператор Робертса. Якщо ми говоримо про виділення сегментів то тут краще підійде ЕМ-метод.

4.5 Перспективи подальших досліджень

Подальші дослідження цієї теми можна продовжувати як окремо, так і в комплексі з іншими завданнями. Наприклад, якщо об'єднати обробку складноструктурованих зображень та роботу з базами даних отримаємо можливість проєкспериментувати над ефективністю пошуку зображень, використовуючи той чи інший метод.

Якщо говоримо про дослідження саме цієї теми, окремо від інших, то треба рухатись в аналіз наборів зображень та їх класифікацію за деякими ознаками. Наприклад, визначення об'єктів, що виділяються від інших: пошкоджені судини, вражені бляшки крові або зіпсовані фрукти чи овочі.

Ця тема може стати значущою за певних умов, таких як зацікавленість бізнесу, збільшення інформації, що добута практичним шляхом, програмне та апаратне забезпечення, необхідність вирішення нових завдань.

ВИСНОВКИ

Протягом дослідження були визначені закономірності серед методів обробки зображень, а також визначити область застосувань дескрипторів Фур'є у сфері обробки даного виду зображень.

Використано методи та основні принципи обробки складноструктурованих зображень, проведений аналіз математичних моделей. Головною задачею даного дослідження є виявлення ефективних гібридних методів для класифікації складноструктурованих зображень.

Проаналізовано методи Канні, ЕМ-метод сегментації, оператори Робертса та Собеля, визначено особливості роботи з бінарними зображеннями, досліджено мережеві та нейромережеві структури для задач класифікації, а також запропоновано методичне та програмне забезпечення, що необхідні для формування вирішальних модулів, використано гібридні методи для класифікації.

ЕМ-метод та метод Канні виділяють сегменти зображення, відповідно до визначених програмою кордонів, а оператори Робертса та Собеля досить чітко визначають кордони об'єктів на зображенні.

Складноструктуроване зображення має містити в собі групу об'єктів різної форми та у різних позиціях (відстань, розмір, кут повороту того чи іншого об'єкта). Саме таке зображення проаналізовано за допомогою вищезазначених методів.

Результати атестаційної роботи апробовано у вигляді тез доповіді на міжнародній конференції.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Творошенко, І. С., & Табашник, В. А. (2018). Розробка просторової моделі геоінформаційної підтримки людей з обмеженими можливостями, що пересуваються на інвалідних колясках, у місті Харків.
2. Творошенко, И. С. (2004). Структура и функции интеллектуальных средств принятия решений в сложных системах. *Искусственный интеллект*, (4), 462–470.
3. Творошенко, И. С. (2010). Анализ процессов принятия решений в интеллектуальных системах. *Системы обработки информации*, (2), 248–253.
3. Кучеренко, Е. И., & Творошенко, И. С. (2010). Прикладные аспекты моделирования нечетких процессов в сложных системах. *Збірник наукових праць Харківського університету Повітряних сил*, (1), 127–131.
5. Кучеренко, Є. І., & Творошенко, І. С. (2011). Оперативне оцінювання простору станів складних розподілених об'єктів з використанням нечіткої інтервальної логіки. *Штучний інтелект*.
6. Творошенко, І. С. (2016). Конспект лекцій з дисципліни «Технології підтримки прийняття рішень в геоінформаційних системах» для студентів 1 курсу денної форми навчання спеціальності 193–Геодезія та землеустрій спеціалізації (освітньої програми)–«Геоінформаційні системи і технології».
7. Кучеренко, Є. І., Кучеренко, В. Є., Глушенкова, І. С., & Творошенко, І. С. (2012). Методи, моделі та інформаційні технології оцінювання станів складних об'єктів: монографія.
8. Кучеренко, Е. И., & Творошенко, И. С. (2003). Процессы принятия решений в сложных системах на основе нечетких интервальных представлений. *Вісник Національного технічного університету "ХПІ". Тематичний випуск: Системний аналіз, управління та інформаційні технології.–Х.: НТУ" ХПІ, 1(7), 79–86.*

9. Кучеренко, Є. І., Творошенко, І. С., & Анопрієнко, Т. В. (2016). Моделювання та оцінювання станів складних об'єктів із застосуванням формальної логіки. *Системи обробки інформації*, (2), 76–82.
10. Кучеренко, Е. И. (2005). Интеллектуальные технологии в задачах принятия решений технологических комплексов на основе нечеткой интервальной логики. *Восточно–Европейский журнал передовых технологий*, (2), 92–96.
11. Мамонов, К. А., & Творошенко, І. С. (2014). Математичні методи і моделі в оцінці нерухомості: навч. посіб.
12. Кучеренко, Е. И., Корниловский, А. В., & Творошенко, И. С. (2010). О методах настройки функций принадлежности в нечетких системах. *Системы управления, навигации и связи*, (1), 13.
13. Бодянский, Е. В., Кучеренко, Е. И., & Творошенко, И. С. (2004). О синтезе нечетких методов на основе композиции фрагментов правил и моделей. *АСУ и приборы автоматизации*, (128), 19–28.
13. Творошенко, І. С. (2015). Конспект лекцій з дисципліни «Цифрова обробка зображень»(для студентів 5 курсу денної та заочної форм навчання спеціальності 7.08010105–Геоінформаційні системи та технології).
15. Творошенко, І. С. Конспект лекцій з дисципліни «Цифрова обробка зображень» для студентів 4 курсу денної форми навчання напряму 6.080101–Геодезія, картографія та землеустрій.
16. Творошенко, І. С. (2018). Основи цифрової обробки зображень: конспект лекцій для студентів 4 курсу денної форми навчання напряму 6.080101–Геодезія, картографія та землеустрій.
17. Творошенко, И. С., & Дехтярь, А. П. (2005, June). Информационные технологии в задачах компьютерной диагностики с использованием интеллектуальных систем. In *Научно–методический журнал Клиническая информатика и телемедицина: Материалы научно–практической конференции с международным участием «Компьютерная медицина 2005»* (Vol. 2, No. 1, p. 138).

18. Творошенко, І. С. (2018). Спеціалізоване програмне забезпечення: конспект лекцій для магістрів денної та заочної форм навчання спеціальності 193–Геодезія та землеустрій освітньої програми «Геодезія та землеустрій».

19. Творошенко, І. С. (2016). Конспект лекцій з дисципліни «Інтелектуальні геоінформаційні системи» для студентів 1 курсу денної форми навчання спеціальності 193–Геодезія та землеустрій спеціалізації (освітньої програми)–«Геоінформаційні системи і технології».

20. Творошенко, І. С., & Шевченко, А. Р. Практичні аспекти застосування геоінформаційного аналізу для забезпечення раціонального благоустрою міста Сєвєродонецька.

21. Творошенко, І. С. (2017). Конспект лекцій з дисципліни «Геоінформаційні системи в управлінні територіями»(для студентів 5 курсу денної форми навчання спеціальностей 7.08010105–Геоінформаційні системи та технології, 8.08010105–Геоінформаційні системи та технології та студентів 6 курсу заочної форми навчання спеціальності 7.08010105–Геоінформаційні системи та технології).

22. Творошенко, І. С. Конспект лекцій з дисципліни «Геоінформаційні системи в задачах моніторингу» для студентів 1 курсу денної форми навчання спеціальності 193–Геодезія та землеустрій спеціалізації (освітньої програми)«Геоінформаційні системи і технології».

23. Творошенко, І. С. (2015). Конспект лекцій з дисципліни «Основи моделювання складних систем»(для студентів 2 курсу заочної форми навчання напряму підготовки 6.080101–Геодезія, картографія та землеустрій).

23. Творошенко, І. С., Мгеброва, В. Р., & Белый, В. В. (2015). Практические аспекты применения современных геоинформационных систем для создания муниципальной геоинформационной системы города Харькова. *Системи обробки інформації*, (7), 65–70.

25. Кучеренко, Є. І., Глушенкова, І. С., & Творошенко, І. С. Методичні вказівки до виконання лабораторних, розрахунково–графічних та самостійних

робіт з дисципліни «Планування та управління ГІС–проектами»(для студентів 5 курсу денної форми навчання зі спеціальностей 7.08010105, 8.08010105 «Геоінформаційні системи і технології» та студентів 6 курсу заочної форми навчання зі спеціальності 7.08010105 «Геоінформаційні системи і технології»).

26. Гончаренко, М. О. (2015). Сравнительный анализ методов формирования дескрипторов изображений в контексте задачи сегментации видеопотока.

27. Гороховатский, В. А. (2014). Структурный анализ и интеллектуальная обработка данных в компьютерном зрении.

28. Putyatin, Y., Gorohovatsky, V., Gorohovatsky, A., & Peredriy, E. (2008). Projective methods of image recognition.

29. Власенко, Н. В. (2013). Построение информативных компактных описаний и классификация объектов путем представления в ортогональном базисе.

30. Путятин, Е. П., Гороховатский, В. А., & Кузьмин, С. В. (2006). Распознавание изображений в пространстве инвариантных локальных признаков. *Радиоэлектроника и информатика*, (1).

31. Куликов, Ю. А., & Лисовин, Д. В. (2010). Исследование структурно-иерархических методов распознавания изображений.

32. Ляшенко, В. В., Кобылин, О. А., & Томич, С. Н. (2016). Основные этапы обработки изображений цитологических препаратов с использованием идеологии вейвлетов.

33. Кочина, М. Л., & Калиманов, В. Г. (2008). Методы обработки изображений для автоматизации диагностики патологии экстракулярных мышц.

34. Гороховатский, В. А. (2008). Применение отношений на множестве характерных признаков изображений при распознавании на основе голосования.

35. Лебедев, В. В., & Селиванова, К. Г. (2019). *Автоматизированная обработка трихоскопических изображений* (Doctoral dissertation, ХНУРЕ).
36. Заворотна, М. Г., & Семенец, В. В. (2019). *Анализ методов и средств реализации алгоритмов распознавания образов по видеоизображению* (Doctoral dissertation, ХНУРЕ).
37. Марченко, Е. Н. (2011). Распознавание и поиск изображений логотипов на основе инвариантных признаков.
38. Богучарский, С. И., & Машталир, С. В. (2014). Кластеризация коллекций изображений в больших базах данных на основе рекуррентной оптимизации.
39. Гороховатский, В. А. (2008). Метрики на множествах ключевых точек изображений.
40. Елаков, С. Г., Руденко, О. Г., & Сотников, О. М. (2012). Предварительная обработка и классификация изображений мелкоструктурных объектов.