

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет навчально-науковий центр заочної форми навчання
(повна назва)

Кафедра електронних обчислювальних машин
(повна назва)

АТЕСТАЦІЙНА РОБОТА
Пояснювальна записка

Рівень вищої освіти другий (магістерський)

Моделі сегментації малорозмірних об'єктів у вбудованих
комп'ютерних системах

(тема)

Виконав:

студент II курсу, групи СПЗм-18-2
Авілова М.С.
(прізвище, ініціали)

Спеціальність

123 – Комп'ютерна інженерія
(код і повна назва спеціальності)

Тип програми освітньо-наукова
(освітньо-професійна або освітньо-наукова)

Освітня програма

Системне програмування
(повна назва освітньої програми)

Керівник: проф. Смеляков К.С.
(посада, прізвище, ініціали)

Допускається до захисту

Зав. кафедри ЕОМ

(підпис)

Коваленко А.А.

(прізвище, ініціали)

2020 р.

Харківський національний університет радіоелектроніки

Факультет _____ навчально-науковий центр заочної форми навчання

Кафедра _____ електронних обчислювальних машин

Рівень вищої освіти _____ другий (магістерський)

Спеціальність _____ 123 – Комп'ютерна інженерія
(код і повна назва)

Тип програми _____ освітньо-наукова
(освітньо-професійна або освітньо-наукова)

Освітня програма _____ Системне програмування
(повна назва)

ЗАТВЕРДЖУЮ:

Зав.

кафедри

_____ (підпис)

“ _____ ” _____ 20__ р.

ЗАВДАННЯ
НА АТЕСТАЦІЙНУ РОБОТУ

студентові _____ Авіловій Марині Сергіївні
(прізвище, ім'я, по батькові)

1. Тема роботи _____ Моделі сегментації малорозмірних об'єктів у вбудованих комп'ютерних системах

затверджена наказом по університету від “ 30 ” березня 2020 р. № 43Стз

2. Термін подання студентом роботи до екзаменаційної комісії _____ 18 травня 2020 р.

3. Вхідні дані до роботи _____

Типи зображень для сегментації

Вимоги до алгоритмів сегментації зображень

Вимоги до застосування згортальних нейронних мереж

4. Перелік питань, що потрібно опрацювати в роботі _____

1) аналіз предметної області

2) огляд основних методів та алгоритмів сегментації зображень

3) аналіз проблематики завдання знаходження об'єктів на зображенні

4) вибір, налаштування, навчання згортальної нейронної мережі

5) Обчислювальний експеримент

6) Висновки

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій (слайдів) _____

Демонстраційні матеріали. Плакати – 14 арк. ф. А4

6. Консультанти розділів роботи (заповнюється 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

№	Назва етапів роботи	Термін виконання етапів роботи	Примітка
1	Аналіз сучасного стану проблеми	31.03.20 – 12.04.20	
2	Постановка задачі дослідження	13.04.20 – 15.04.20	
3	Аналіз літератури	16.04.20 – 20.04.20	
4	Розгляд методів та алгоритмів сегментації	21.04.20 – 24.04.20	
5	Аналіз сучасних рішень сегментації	25.04.20 – 12.05.20	
6	Оформлення пояснювальної записки	13.05.2020 - 14.05.20	
7	Подання атестаційної роботи керівникові	12.05.20 – 14.05.20	
8	Рецензування атестаційної роботи	15.05.20 – 17.05.20	
9	Подання атестаційної роботи до ЕК	18.05.20	

Дата видачі завдання 30 березня 2020 р.

Студент _____
(підпис)

Керівник роботи _____
(підпис)

проф. Смеляков К.С.
(посада, прізвище, ініціали)

РЕФЕРАТ

Пояснювальна записка атестаційної роботи: 87 с., 37 рис., 6 табл., 1 додаток, 12 джерел.

СЕГМЕНТАЦІЯ, ЗГОРТАЛЬНА НЕЙРОННА МЕРЕЖА, СТРОБ, СТРУКТУРНИЙ ЕЛЕМЕНТ, МАЛОРОЗМІРНІ ЗОБРАЖЕННЯ, АЛГОРИМ СУПРОВОДУ, МОДЕЛЬ.

Метою атестаційної роботи є дослідити процес сегментації малорозмірних об'єктів у вбудованих комп'ютерних системах.

У ході виконання атестаційної роботи був проведений детальний аналіз предметної області. Були розглянуті методи сегментації зображень, які використовуються у різних галузях сучасного життя. Для предмету аналізу було обрано ряд алгоритмів та були сформовані критерії оцінки якості для кожного.

Проведений аналіз результатів експериментів методів сегментації у різних умовах. Виявлені їх переваги та недоліки.

ABSTRACT

Master's thesis: 87 pages, 37 figures, 6 tables, 1 appendix, 12 sources.

SEGMENTATION, CONVOLUTIONAL NEURAL NETWORK, STROB, STRUCTURAL ELEMENT, SMALL IMAGES, ALGORIM OF SUPPORT, ALGORIM OF SUPPORT, MODEL.

The purpose of certification work is to investigate the process of small objects segmentation in embedded computer systems.

During the certification work, a detailed analysis of the subject area was conducted.

Methods of image segmentation used in various areas of modern life were considered. A number of algorithms were selected for the subject of the analysis and quality assessment criteria were formed for each.

The analysis of the models segmentation experiments in different conditions is carried out. Their advantages and disadvantages are revealed.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ	8
ВСТУП	9
1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ	10
1.1 Виявлення об'єктів для супроводу	11
1.1.1 Алгоритм супроводу	13
1.1.2 Виділення меж Кані	16
1.1.3 Математична морфологія	18
1.1.4 Модель і критерії сегментації МЗ	26
1.2 Виділення рухомих об'єктів у відеопотоці	30
1.3 Згортальні нейронні мережі	31
1.3.1 Архітектура і принцип роботи згортальної нейронної мережі	31
1.3.2 Робота згортальної нейронної мережі на прикладі сегментації сітки сітківки ока за допомоги нейромережі (U-net)	34
1.3.3 Оцінка навченої моделі	37
2 МЕТОДИ ВИРІШЕННЯ ЗАДАЧІ СЕГМЕНТАЦІЇ МАЛОРОЗМІРНИХ ОБ'ЄКТІВ У РІЗНИХ УМОВАХ	40
2.1 Метод автоматичного виявлення та вистежування багатьох малорозмірних об'єктів в умовах в умовах апріорної невизначеності	40
2.1.1 Опис роботи методу	42
2.1.2 Стробовий метод виявлення малорозмірних об'єктів	44
2.1.3 Дослідження ефективності методу	47
2.1.4 Підсумок за методом автоматичного виявлення та вистежування багатьох малорозмірних об'єктів в умовах в умовах апріорної невизначеності	49
2.2 Ідентифікаційний метод розпізнавання образів малорозмірних підводних об'єктів	49

2.2.1	Задачі що вирішує ідентифікаційний метод	50
2.2.2	Результати експерименту	56
2.2.3	Підсумок за ідентифікаційним методом розпізнавання образів малорозмірних підводних об'єктів.....	64
3 АНАЛІЗ ТА КЛАСИФІКАЦІЯ АЛГОРИТМІВ СЕГМЕНТАЦІЇ ЗОБРАЖЕНЬ		66
3.1	Узагальнена класифікація алгоритмів сегментації зображень.....	67
3.2	Алгоритм квазіоптимальний сегментації.	69
3.3	Класифікації за п'ятьма ознаками.....	73
3.4	Аналіз класифікаційних ознак	75
ВИСНОВКИ.....		78
ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ		79
ДОДАТОК А ГРАФІЧНИЙ МАТЕРІАЛ АТЕСТАЦІЙНОЇ РОБОТИ		81

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ

ВСШ – відносини сигнал-шум

ІА - ідентифікаційного аналізатора

РМЗ - розподіл миттєвих значень

ІАЦ – інформаційно-аналітичний центр

ГБЗ - гідролокатори бокового огляду

МЗ – малорозмірні зображення

ФАР - фазована антенна решітка

ВСШ - відносна сигнал-шум

CNN - Convolutional neural network

TBD - Track Before Detect

АСЗ - алгоритм сегментації зображень

ВСТУП

Розробкою алгоритмів знаходження в автоматичному режимі малорозмірних об'єктів на растрових зображеннях фахівці з цифрової обробки зображень займаються вже не одне десятиліття. Під малорозмірним об'єктом в даному випадку розуміється зображення об'єкта, що займає кілька десятків пікселів на зображенні сцени. Результатом роботи алгоритму знаходження малорозмірного об'єкта вважається виділення на зображенні групи пікселів, що відносяться до шуканого об'єкту. Незважаючи на досягнуті успіхи в рішенні окремих завдань, проблема все ще залишається актуальною. Конкретні алгоритми істотно залежать як від мети виявлення малорозмірних об'єктів, так і від обмежень на зміст сцени.

Метою виявлення об'єкта на зображенні можуть бути, наприклад, тільки його розпізнавання, тобто віднесення до одного з відомих класів, наприклад, розпізнавання друкованих символів, або підрахунок кількості об'єктів, що відносяться до різних класів, наприклад, збір статистичних даних з використанням аерофотознімків, або супровід виділених як рухомих, так і нерухомих об'єктів по кадрах відеопотоку, наприклад, для моніторингу транспортних потоків, чи визначення положення виявленого і розпізнаного об'єкта, наприклад, для суміщення зображень, отриманих з різних ракурсів.

Обмеження на зміст сцени стосуються як власне виділяються об'єктів, наприклад, можливих символів при обробці растрових зображень тексту, так і фону, на якому необхідно виявити об'єкти. Фон на зображенні з малорозмірними об'єктами може бути як рівномірним, наприклад, при виявленні повітряних об'єктів, так і мати складну структуру, наприклад, в системах відеоспостереження в міському середовищі.

У даній роботі докладніше розглянемо найбільш характерні з існуючих в даний час алгоритмів пошуку малорозмірних об'єктів.

1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ

1.1 Виявлення об'єктів для супроводу

Для знаходження малорозмірних об'єктів на зображеннях з рівномірним фоном (море, небо, поле) застосовуються кореляційні методи [5] як в просторовій області, так і в частотній області з використанням різних ортогональних перетворень (Фур'є, Уолша, Адамара).

Алгоритми, розроблені для виявлення і супроводу об'єктів, найбільш ефективні при наявності нерухомого фону з переміщається по ньому об'єктом. Положення об'єкта на зображенні визначається координатами пікселя, відповідного центру описаного навколо зображення об'єкта прямокутника. У разі повітряних об'єктів фон являє собою практично постійну або плавно змінюється функцію яскравості з включеннями хмар, швидкість переміщення яких значно менше швидкості виділяється об'єкта. Для оцінки параметрів авторегрессионной моделі фону з метою його подальшої компенсації використовуються як алгоритми лінійного передбачення, так і адаптивної просторової фільтрації. Результати порівняльного аналізу показують, що адаптивний алгоритм цілком задовільно працює при відношенні сигнал / шум 2,5 і більше.

Для зображень штучних об'єктів характерна наявність прямолінійних кордонів (будівлі, дороги). Зображення природних об'єктів в основному мають криволінійні кордону, хоча іноді можуть зустрічатися і прямі лінії, наприклад, лінія горизонту. Межі об'єктів на зображенні є області перепаду яскравості і для їх виділення застосовуються добре відомі диференціюються або різницеві фільтри. Наприклад, LoG-фільтр, що дозволяє виділити лінію контуру товщиною в один піксель з одночасним видаленням шуму і дрібних малоконтрастних деталей.

Для виділення тільки прямолінійних кордонів застосовуються більш складні методи, наприклад, виділення ліній перепаду яскравості товщиною в один піксель детектором Кані з подальшим застосуванням перетворення Хафа. Недолік такого підходу полягає в тому, що детектор Канні, як відомо, через нелінійність оператора немаксимального придушення дає на виході «тремтячу» лінію, а також згладжує кути, тобто точки різкої зміни напрямку кордону. Перетворення Хафа в цьому випадку дає фрагментовану лінію, що вимагає додаткової обробки для зв'язування фрагментів, а для виділення кутових точок доводиться застосовувати додаткові алгоритми, наприклад, детектор кутів Харріса.

Нещодавно запропонований метод, що дозволяє поліпшити результати виділення прямих ліній на зображенні, заснований на виконанні орієнтованої фільтрації з подальшим пошуком прямолінійних сегментів граничних ліній по кожному напрямку з різними коефіцієнтами масштабу. Знайдені прямолінійні сегменти упорядковуються у напрямку і використовуються для побудови ієрархічної системи дескрипторів розпізнаваних об'єктів, зображення яких обмежені прямими лініями. Виділення початкових і кінцевих точок прямих ліній проводиться після формування профілю градієнта в обраному напрямку.

Бінаризація зображення виконується з урахуванням результатів ерозії ізольованих фрагментів, що покращує виділення протяжних об'єктів на фоні перешкод різного виду. Поріг бінаризації вибирається з урахуванням результатів сегментації на основі ієрархії за формою і розмірами ізольованих фрагментів, що утворюються на бінарному зображенні в результаті фільтрації набором відповідних масок. Експерименти, проведені розробниками даного методу, показали високу ефективність методу для виділення і опису об'єктів штучного походження на аерофотознімки.

Виділення на зображенні об'єктів природного походження представляє значні труднощі в силу наявності, як правило, криволінійної кордону, яка відділяє об'єкти як один від одного, так і від фону. Значною мірою завдання

полегшується введенням додаткових обмежень, що визначають умови отримання зображення або специфіку, що виділяються. Наприклад, розроблені досить ефективні методи пошуку на аерофотознімки стадних тварин з метою підрахунку їх поголів'я.

В даному випадку автоматична система повинна підраховувати на знімку загальне число тварин, які можуть являти собою як окремі особини, так і нерозділені в процесі сегментації групи, причому на знімку можлива наявність декількох видів візуально помітних тварин.

Специфічні особливості предмета обробки полягають в наступному. По-перше, система призначена для підрахунку поголів'я північних оленів з використанням аерофотознімків тундри, що забезпечує практично рівномірний фон на зображенні. По-друге, цілком допустимим є деякий відсоток помилок при розпізнаванні, тому що результатом роботи системи є статистичні дані. Зазначені особливості дозволили, використовуючи добре відомі методи попередньої обробки, бинаризації, сегментації та розпізнавання зображень за геометричними характеристиками сегментів, отримати цілком задовільні результати.

Недоліки системи полягають в тому, що при зміні колірного балансу знімка результати сегментації змінюються, а головне - при підрахунку враховуються сторонні об'єкти, що мають геометричні характеристики, схожі з характеристиками розпізнаваних об'єктів.

При дистанційному зондуванні земної поверхні за допомогою многополосних спектрометрів часто потрібно не тільки розпізнати і виділити на зображенні деяку область, а й визначити її справжні розміри. Для цього необхідно виконати корекцію просторових спотворень, обумовлену різними ракурсами зйомки об'єкта. При корекції просторових спотворень плоских об'єктів для знаходження параметрів афінного перетворення використовуються координати характерних точок.

У системах технічного зору при суміщенні растрових зображень, отриманих з різних ракурсів, для виключення просторових спотворень

виконується прив'язка реперних точок, за координатами яких обчислюються параметри перетворення суміщаються зображень.

1.1.1 Алгоритм супроводу

У проблемах, що підтримують об'єкти на складному телефоні телевізійними методами, виведені наступні завдання:

- об'єкти первинної сегментації;
- співпрацювання об'єкта в умовах динаміки свого власного, а також взаємодія з іншими об'єктами та фоном.

У одночасному кадрі видовища сегментації об'єктів реалізують за визнанням яркості, локальної детальності (з високою потужністю енергії) і форм. Аналіз відеопослідовності дозволяє використовувати визнання руху. Зчитування розпізнавання руху забезпечує надійну сегментацію обласного зображення, надає можливість переміщення об'єктів на неподвижному, в тонкому стильному, тонкому кольорі.

Традиційно для оцінювання визнають, що рухаються, застосовуючи енергетичні рухи. Під енергетичними рухами відображаються знімки яркості пікселів у сусідніх кадрах відеопослідовності (виробнича яркостна сигнала за часом)

$$\frac{dL(x, y, t)}{dt} \approx \frac{L(x, y, t) - L(x, y, t - \Delta t)}{\Delta t}, \quad (1.1)$$

де $L(x, y, V)$ - яркість пікселя з координатами (x, y) в момент часу t .

Скалярна оцінка у вигляді енергії, що рухається, не забезпечує різних об'єктів, що знаходяться в безпосередній близькості до іншого і рухається з різними скоростями. Более інформативно поле векторів руху (оптичний потік). Після проектування на плоскості кадра реальне трохмерне переміщення об'єктів показується двомірним переміщенням, яке може бути

оцінено миттєвими швидкостями або дискретними змінами фрагментів зображення - полем векторів руху. Методи на основі визнання переміщення неможливо вивести стаціонарні об'єкти. Визначені важкості випромінюють сегментацію дуже повільно переміщуються об'єктів (видима швидкість руху менше одного пікселя). Одночасне використання поля векторів переміщення, а також визнання просторової свідомості та детальності можливо сегментувати декілька об'єктів, у тому чисті неподільні, на неоднородному (складному) кольорі.

На підставі первинної сегментації формуються моделі об'єктів, що цікавлять, для яких визначають найкраще відповідність у наступних кадрах. У кореляційно-екстремальних системах використовуються моделі об'єктів, що відрізняються різними степенями складності (деталізаціями) в залежності від априорної інформації про об'єкти. За найвищим алфавітом об'єктів, можливо, використовували 3D-комп'ютерні моделі з наступною підготовкою бінарної маски або контурного препарату (рис. 1.1). При незвіданому алфавіті моделі строго базуються на основі аналізу сегментованої області зображення. Автоматичне побудовані моделі ефективно реалізують на головному тоні. На складному телефоні, необхідному для участі оператора, який формує строб, охоплюючий об'єкт. У випадку, коли всі зображення всередині строби розглядаються як модель об'єкта. Для локалізації власних об'єктів всередині вікна виходить контурний препарат за допомогою операторів - вертикальних, горизонтальних та нахилених контурів.



Рисунок 1.1 – 3D модель літального апарату, побудована за сеансами та бінарною маскою

Процедура найбільш успішна при відсутності контурів на фрагментах фона, попавших у строб. Зчитуйте, що в кореляційних системах основними проблемами використовуються втрати об'єкта, що цікавлять на складному тоні в умовах динаміки власних об'єктів, а також в момент перекриття об'єкта цікавих інших об'єктів або фрагментів фона.

Об'єднання визнають деталі та форми із визнаними переміщеннями на етапі співпрацювання можливо реалізувати слоєве представлення: кожен об'єкт об'єкта та розміщує свій шар. Це забезпечує автоматичне об'єднання об'єктів на складному тоні, вирішення ситуацій оключення, сходження та розділення об'єктів. Сегментація та підтримка об'єктів на основі їх слоєвого представлення основні процедури, запропоновані методом використання:

Сегментація і супровід об'єктів на основі їх шарового подання.

Основними процедурами запропонованого методу є:

- попередній аналіз зображення на основі визнання детальності;
- сегментація об'єктів, що цікавляться на основі визнання просторової свідомості та руху;
- співпрацювання об'єктів на основі визнаних рухів, зв'язків і форм (визнають, що навчають з адаптивним весом, рис. 1.2).

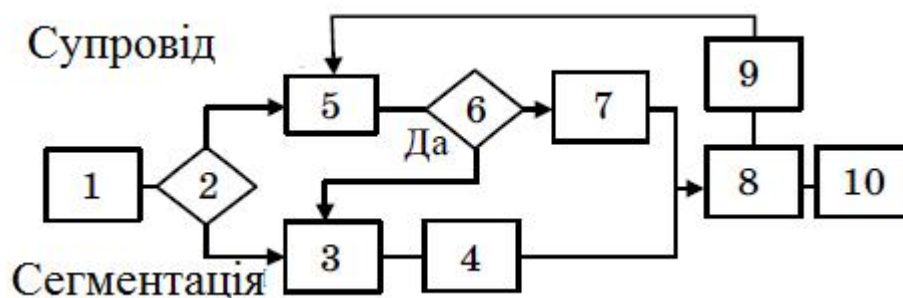


Рисунок 1.2 – Узагальнена структура сегментації і супроводу об'єктів на основі їх шарового уявлення

Взаємодія перехідних процедур ілюструється на рис. 1.2, де:

- 1 - попередня класифікація зображень на основі визнання детальності;

- 2 - вид режиму: первинна сегментація / сопровождение;
- 3 - сегментація об'єктів на основі просторової свідомості та руху;
- 4 - формування моделей форм і переміщення сегментованих об'єктів;
- 5 - співпрацювання об'єктів на основі визнаних рухів і форм, навчальних з адаптивних вісами;
- 6 - перевірка: є блокована, не віднесена до більш раннього сегментованого об'єкта інтересів;
- 7 - оновлення моделей форм і переміщення з участю предказанія;
- 8 - аналіз власності об'єкта та окружної реставрації;
- 9 - адаптація весів визначників;
- 10 - представлення отриманої інформації про об'єкти та вироблення керуючих сигналів.

1.1.2 Виділення меж Кані

Кані (John F. Canny; 1953 г.) вивчив математичну проблему отримання фільтра, оптимального за критеріями виділення, локалізації та мінімізації декількох відгуків одного краю. Це означає, що детектор повинен реагувати на межі, але при цьому ігнорувати помилкові, точно визначати лінію кордону (без її фрагментованість) і реагувати на кожну кордон один раз, що дозволяє уникнути сприйняття широких смуг зміни яскравості як сукупності кордонів. Канні ввів поняття Non-Maximum Suppression (придушення НЕ-максимумів), яке означає, що пікселями кордонів оголошуються точки, в яких досягається локальний максимум градієнта в напрямку вектора градієнта.

Хоча його робота була проведена на зорі комп'ютерного зору (1986 рік), детектор кордонів Канні досі є одним з кращих детекторів. Крім особливих окремих випадків важко знайти детектор, який би працював істотно краще, ніж детектор Кані. Завдання: демонстрації роботи алгоритму буде вирішена задача по виділенню кордонів на наступному зображенні.

Алгоритм складається з п'яти окремих кроків:

- згладжування. Розумієте зображення для видалення шуму.
- пошук градієнтів. Межі відзначаються там, де градієнт зображення набуває максимальне значення.
- придушення НЕ-максимумів. Тільки локальні максимуми відзначаються як кордони.
- подвійна порогова фільтрація. Потенційні межі визначаються порогоми.
- трасування області неоднозначності. Підсумкові межі визначаються шляхом придушення всіх країв, незв'язаних з певними (сильними) межами.

Перед застосуванням детектора, перетворимо зображення в відтінки сірого, щоб зменшити обчислювальні витрати. Цей етап характерний для багатьох методів обробки зображень (рис. 1.3).

Для перетворення вихідного зображення в зображення в градаціях сірого, необхідно отримати його «яскравість» -складати. Для цього зручно представити зображення в колірній моделі YUV (або HSL, HSV, інших).

Спочатку малюнок представлений в RGB колірної моделі, причому функція завантаження повертає три компонента у вигляді однієї матриці.



а)

б)

Рисунок 1.3 – Зображення: а) вихідне зображення і б) результат обробки.

1.1.3 Математична морфологія

Термін морфологія відноситься до опису властивостей форми і структури будь-яких об'єктів. В контексті машинного зору цей термін

відноситься до опису властивостей форми областей на зображенні. Операції математичної морфології спочатку були визначені як операції над множинами, але скоро з'ясувалося, що вони також корисні в задачах обробки безлічі точок в двовимірному просторі. Множинами в математичній морфології представляються об'єкти на зображенні.

В першу чергу математична морфологія використовується для вилучення деяких властивостей зображення, корисних для його уявлення і опису. Наприклад, контурів, кістяків, опуклих оболонок. Також інтерес представляють морфологічні методи, що застосовуються на етапах попередньої та підсумкової обробки зображень. Наприклад, морфологічна фільтрація, потовщення або утоньшення.

Вхідними даними для апарату математичної морфології є два зображення: обробляється і спеціальне, залежне від виду операції та розв'язуваної задачі. Таке спеціальне зображення прийнято називати примітивом чи структурним елементом. Структурний елемент можна вважати опис області з деякою формою. Зрозуміло, що форма може бути будь-який, головне, щоб її можна було уявити у вигляді бінарного зображення заданого розміру. У багатьох пакетах обробки зображень найбільш поширені структурні елементи мають спеціальні назви: BOX [H, W] -прямоугольник заданого розміру, DISK [R] - диск заданого розміру, RING [R] - кільце заданого розміру (рис. 1.4).

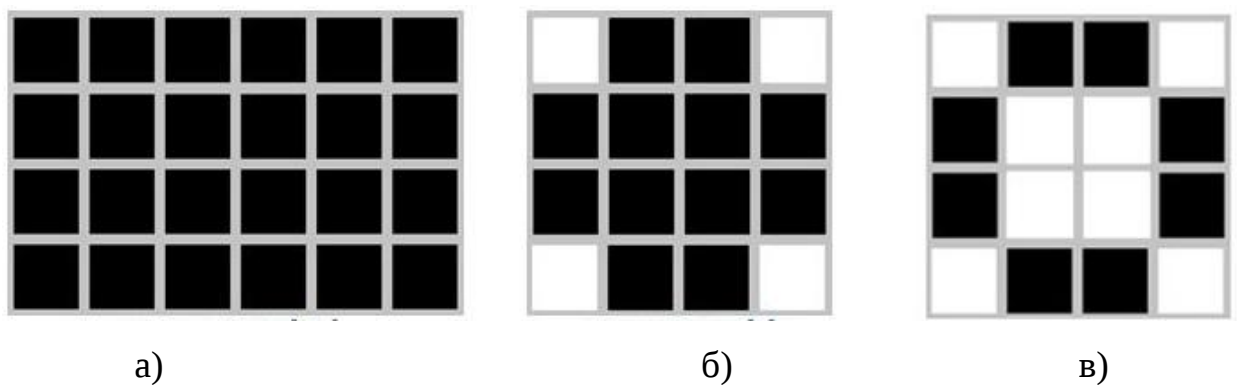


Рисунок 1.4 – Шаблони: а) BOX [4,6], б) DISK [4], в) ING [4].

Результат морфологічної обробки залежить як від розміру і конфігурації вихідного зображення, так і від структурного примітиву.

Розмір структурного елементу як правило дорівнює 3×3 , 4×4 або 5×5 пікселів. Це обумовлено головною ідеєю морфологічної обробки, в процесі якої відшукуються характерні деталі зображення. Шукана деталь описується примітивом, і в результаті морфологічної обробки можна підкреслити або видалити такі деталі на всьому зображенні.

Одне з основних переваг морфологічної обробки - її простота: як на вході, так і на виході процедури обробки ми отримуємо бінаризованими зображення. Інші методи, як правило, з вихідного зображення спочатку отримують півтонове, яке потім приводиться до бінарним за допомогою порогової функції.

Основними операціями математичної морфології є нарощування, ерозія, замикання і розмикання. У цих назвах відображена суть операцій: нарощування збільшує область зображення, а ерозія робить її менше, операція замикання дозволяє замкнути внутрішні отвори області та усунути затоки вздовж кордону області, операція розмикання допомагає позбутися від маленьких фрагментів, які виступають назовні області поблизу її кордону. Далі будуть представлені математичні визначення морфологічних операцій (об'єднання, перетин, доповнення, різниця).

Перш, ніж переходити до операцій морфології, є сенс розглянути теоретико-множинні операції, що лежать в основі математичної морфології.

Об'єднання двох множин A і B , яке позначається $C = A \cup B$, є за визначенням безліч всіх елементів, що належать або множині A , або множині B , або обом множинам одночасно. Аналогічно, перетин двох множин A і B , яке позначається $C = A \cap B$, є за визначенням безліч всіх елементів, що належать одночасно обом множинам A і B . Доповнення множини A є безліч елементів, які не містяться в A : $A^c = \{w \mid w \notin A\}$. Різниця двох множин A і B позначається $A \setminus B$ і визначається наступним чином: $A \setminus B = \{w \mid w \in A, w \notin B\} = A \cap B^c$. Це безліч складається з елементів A , які не входять до безліч B .

Розглянемо всі перераховані вище операції на конкретному прикладі (рис. 1.5).

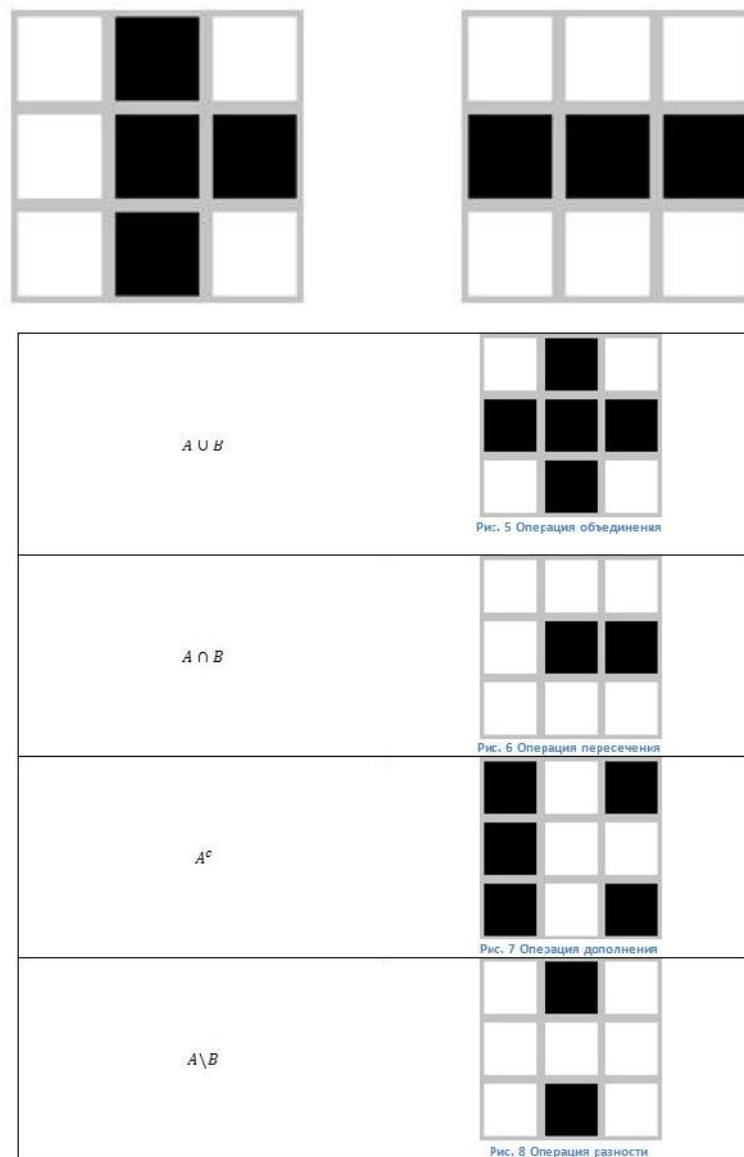


Рисунок 1.5 – Зображення A и B

Перенесення. Операція перенесення Xt безлічі пікселів X на вектор t задається в вигляді $Xt = \{x + t \mid x \in X\}$. Отже, перенесення безлічі одиничних пікселів на бінарному зображенні зрушує все пікселі безлічі на задану відстань. Вектор перенесення t може здаватися у вигляді впорядкованої пари $(\Delta r, \Delta c)$, де Δr - компонент вектора перенесення в напрямку рядків, а Δc - компонент вектора перенесення в напрямку стовпців зображення (рис. 1.6).

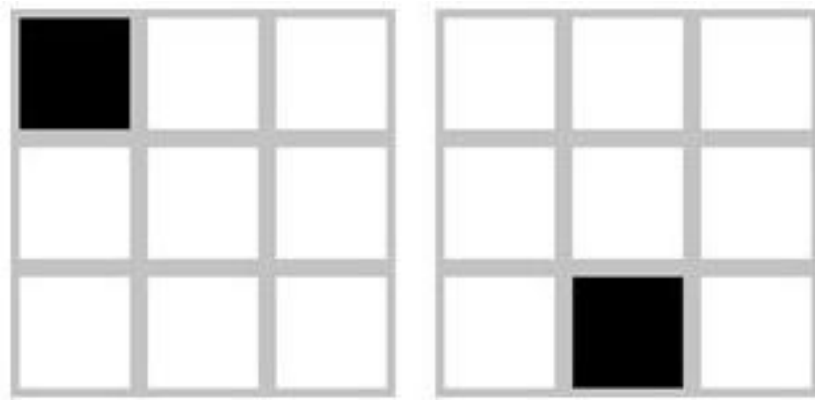


Рисунок 1.6 – Застосування перенесення при $t=(2,1)$

Нарощування. Структурний елемент S застосовується до всіх пікселам бінарного зображення. Кожен раз, коли початок координат структурного елементу поєднується з одиничним бінарним пікселом, до всього структурному елементу застосовується перенесення і подальше роз'єднання з відповідними пікселями бінарного зображення. Результати логічного складання записуються в вихідний бінарне зображення, яке спочатку ініціалізується нульовими значеннями (рис. 1.7, рис. 1.8).

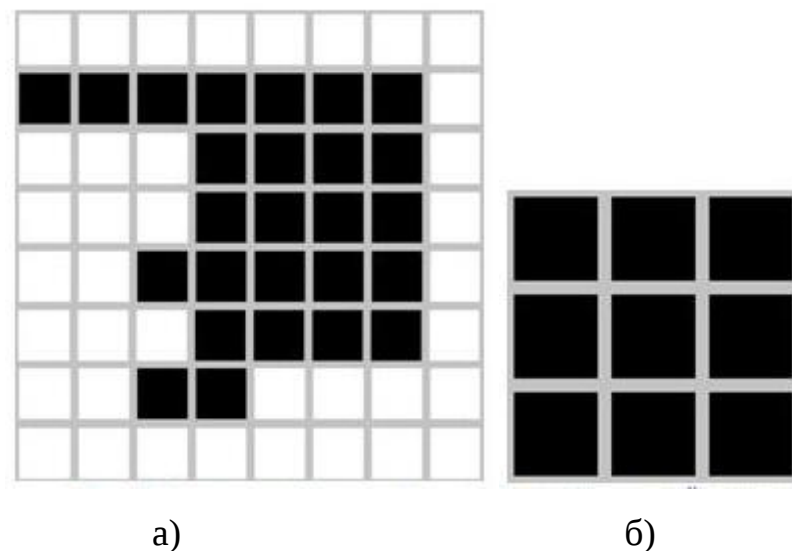


Рисунок 1.7 – Обробка елементом: а) бінарне зображення, б) структурний елемент S

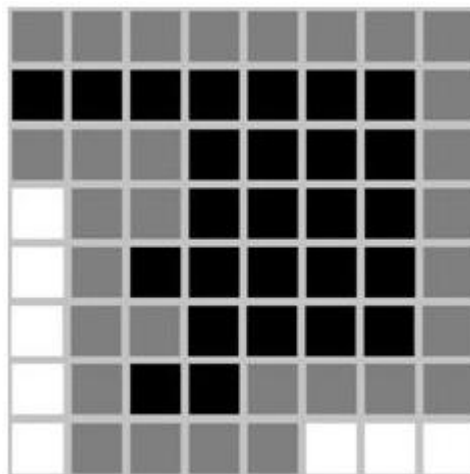


Рисунок 1.8 – Нарощування зображення В структурним елементом S^2

Ерозія. При виконанні операції ерозії структурний елемент теж проходить по всім пікселям зображення. Якщо в деякій позиції кожен одиничний піксель структурного елементу співпадає з одиничним пікселем бінарного зображення, то виконується роз'єднання центрального пікселя структурного елементу з відповідним пікселем вихідного зображення.

В результаті застосування операції ерозії всі об'єкти, менші ніж структурний елемент, стираються, об'єкти, з'єднані тонкими лініями стають роз'єднаними і розміри всіх об'єктів зменшуються (рис. 1.9).

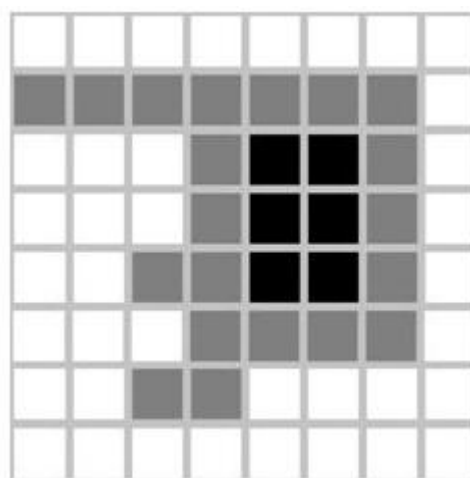


Рисунок 1.9 – Ерозія зображення В структурним елементом S^2 .

Розмикання. Операція ерозії корисна для видалення малих об'єктів і різних шумів, але у цій операції є недолік - всі, хто залишився об'єкти зменшуються в розмірі. Цього ефекту можна уникнути, якщо після операції ерозії застосувати операцію нарощування з тим же структурним елементом.

Розмикання відсіває всі об'єкти, менші ніж структурний елемент, але при цьому допомагає уникнути сильного зменшення розміру об'єктів. Також розмикання ідеально підходить для видалення ліній, товщина яких менше, ніж діаметр структурного елементу. Також важливо пам'ятати, що після цієї операції контури об'єктів стають більш гладкими (рис. 1.10, рис. 1.11).

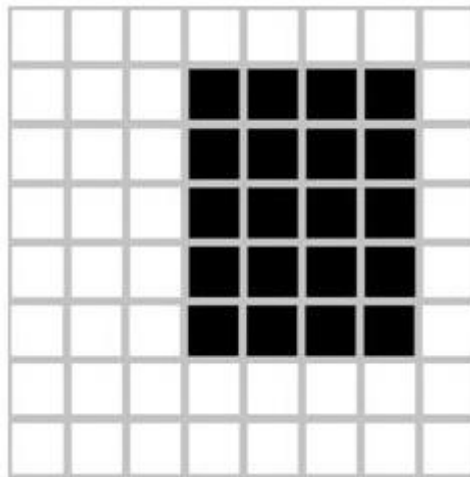


Рисунок 1.10 – Розмикання зображення В структурним елементом S

Зімканення. Якщо до зображення застосувати спочатку операцію нарощування, то ми зможемо позбутися від малих дірок і щілин, але при цьому відбудеться збільшення контуру об'єкта. Уникнути цього збільшення дозволяє операція ерозія, виконана відразу після нарощування з тим же структурним елементом.

Умове нарощування. Одним з типових застосувань бінарної морфології є виділення на бінарному зображенні компонент, у яких форма і розміри задовольняють заданим обмеженням.

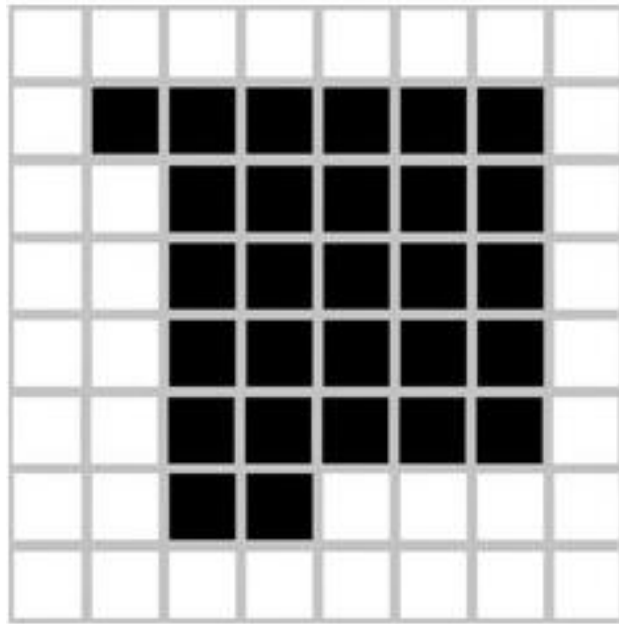


Рисунок 1.11 – Зімкнення зображення В структурним елементом S

У багатьох подібних завданнях можлива побудова структурного елемента, який після застосування до бінарним зображенню видаляє не задовольняють обмеженням компоненти і залишає кілька одиничних пікселів, що відповідають задовольняє обмеженням компонентів.

Але для подальшої обробки можуть знадобитися компоненти цілком, а не тільки їх фрагменти, що залишилися після ерозії. Для вирішення цієї проблеми була введена операція умовного нарощування.

Безліч отримане в результаті ерозії циклічно нарощується структурним елементом S, і на кожному кроці результат зменшується до підмножини пікселів, які мають одиничні значення на оригінальному документі В. Операція умовного нарощування пояснена на малюнку нижче.

На цьому малюнку бінарне зображення В було піддано ерозії елементом V для виділення компонент, що містять вертикальні фрагменти висотою 3 пікселя.

На отриманому зображенні С є дві таких компоненти. Щоб виділити ці компоненти цілком, зображення С умовно нарощується елементом D щодо вихідного зображення В (рис. 1.12).

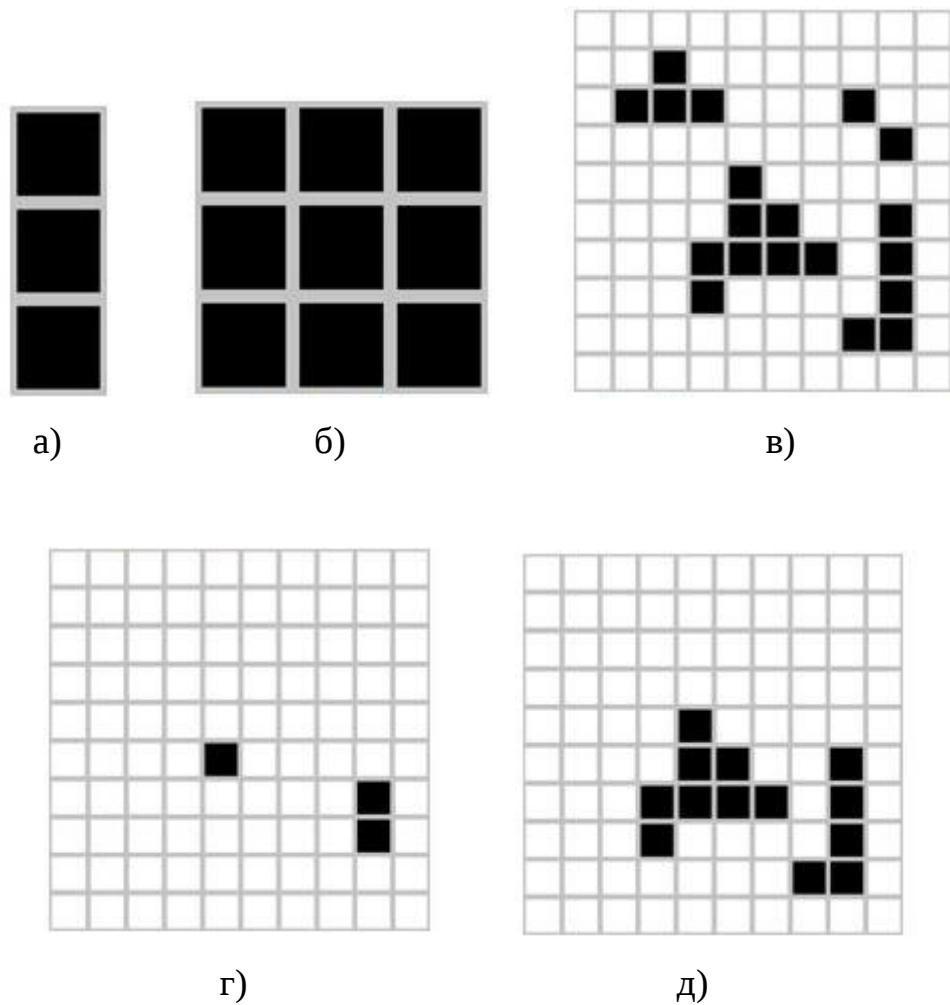


Рисунок 1.12 – Умовне нарощування, де: а) Структурнтй елемент V, б) Структурнтй елемент D, в) Бінарне зображення B, г) Ерозія зображення B структурним елементом V, д) Умовне нарощування заборазення структурним елементом D

Виділення меж. Морфологічні операції можна також використовувати для виділення меж бінарного об'єкта. Це операція дуже важлива, тому що межа є повним, і в той же час досить компактным описом об'єкта (рис. 1.13).

Легко помітити, що граничні точки мають як мінімум один фоновий піксель в своїй околиці. Таким чином, застосувавши оператор ерозії з структурним елементом, що містить всі можливі сусідні елементи, ми видалимо всі граничні точки. Тоді межа вийде за допомогою операції різниці множин між вихідним зображенням і тим, що отримане в результаті ерозії.

належать множині типових відстаней $\{\rho_\zeta\}_{\zeta=1,2,\dots} = \{1, \sqrt{2}, 2, \sqrt{5}, \sqrt{8}, 3, \dots\}$ між пікселями дискретного поля знімка.

Одновимірний випадок. При аналізі щодо виділеного напрямку контрастність пікселя d_i з яскравістю f_i з використанням одновимірної маски радіусу ρ_ζ щодо пікселя d_j з яскравістю f_j оцінюється так

$$u_{i,\zeta} = |f_i - f_j|. \quad (1.2)$$

Використання одновимірних масок служить для оперативного виявлення МЗ на початковій фазі сегментації з метою локалізації області пошуку МЗ. При цьому збільшення радіусу маски дозволяє оцінювати контрастність пікселів МЗ щодо фону, а не тіні, яскравість пікселів якої значимо варіюється; знаючи, що яскравість МЗ більше (або менше) яскравості фону, контрастні пікселі фону можна відфільтрувати за знаком виразу $(f_i - f_j)$ в (1.2). Під тінню зображення тут розуміється безліч пікселів вхідного зображення, яскравість яких після оцифровки утворена випромінюванням об'єкта і фону (рис. 1.14).

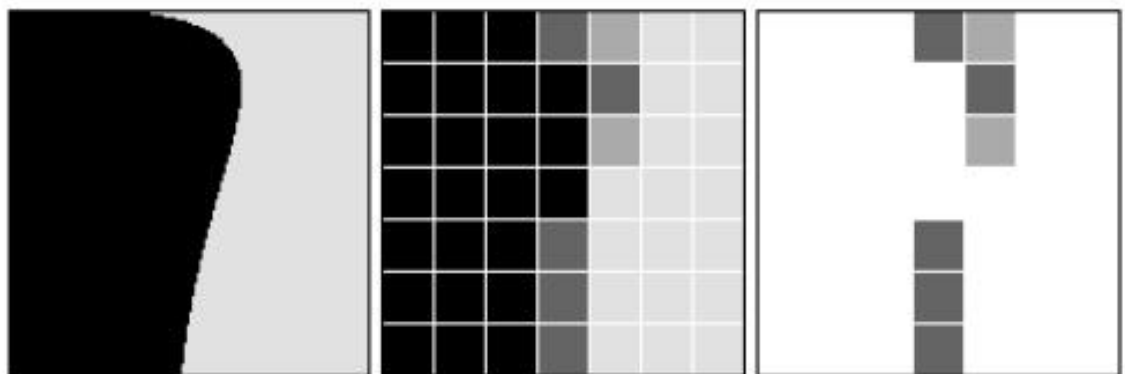


Рисунок 1.14 – Зображення: а) оптичне зображення об'єкта, б) цифрове зображення, в) Тінь об'єкта

Двовимірний випадок. Нехай задано однозв'язного безліч G - регіон. Його огинає $l_{\zeta,\eta}$ - це безліч пікселів, розташованих на відстані не ближче ρ_ζ і не далі ρ_η від кордону регіону G . Сукупність регіону і його обвідної назовемо регіональною маскою $mask_{\zeta,\eta}$ (рис. 1.15).

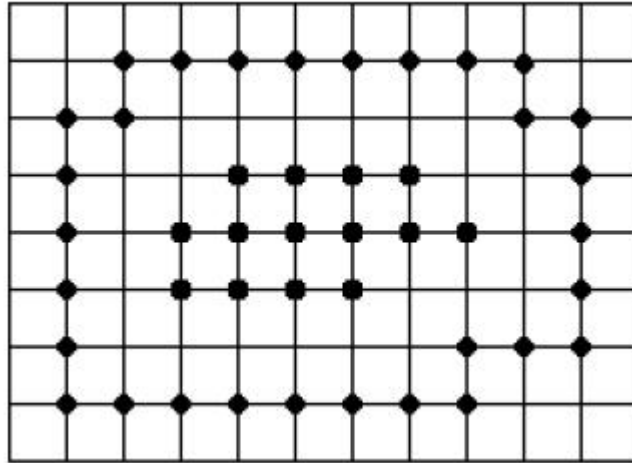


Рисунок 1.15 – Регіональна маска $mask_{3,5}$, де чорним відмічені пікселі регіона G , а сірим – пікселі $l_{3,5}$ що огинає

Нехай f_G і f_l - середні яскравості регіону і його обвідної. Тоді контрастність регіону G оцінимо так

$$\bar{u}_G = |\bar{f}_G - \bar{f}_l|. \quad (1.3)$$

Ці моделі далі використовуються для опису властивостей МЗ з метою його сегментації по контрастності за допомогою адаптації форми регіону на форму МЗ. При цьому можливість зміни радіусів ρ_ξ і ρ_η дозволяє домагатися ситуації, коли при суміщенні МЗ і регіону G його огинає складається з пікселів фону, а тінь міститься між регіоном і його обвідної, що потрібно для отделимости МЗ від фону. Таким чином, основні елементи моделі МІ, на додаток до обмежень (1.3), які служать для швидкої ідентифікації об'єктів по діаметру, представлені: радіусами масок ρ_ζ , ρ_ξ , ρ_η , нижньої оцінкою

контрастності (порогом) T_1 для швидкого виявлення МІ з використанням одновимірної маски по (1.2), а також нижньої оцінкою контрастності (порогом) T_2 для виділення МЗ з використанням регіональної маски по (1.3). Крім цього, для зниження ступеня маскування помилкових пікселів в процесі сегментації може перевірятися виконання обмеження на неоднорідність яркостей пікселів одержуваного регіону. Критерій Kr_1 сегментації пікселя МЗ: піксель вхідного зображення d_i вважається пікселем шуканого МЗ, якщо для нього виконуються співвідношення $u_{i;\zeta} \geq T_1$, за умови, що $f_i > f_j$ (припускаємо, що яскравість фону нижче яскравості об'єкта). Критерій Kr_2 сегментації МЗ: регіон G вважається потрібним МЗ, якщо його діаметр адовольняє (1.1), а контрастність задовольняє співвідношенню $u_G \geq T_2$, при умові, що $f_G > f_l$.

Сегментація МЗ. На основі введеної системи критеріїв пропонується метод адаптивної сегментації МІ, основні кроки якого такі.

Крок 1. За критерієм Kr_1 порядковим скануванням вхідного зображення шукаємо контрастний піксель; при виявленні такого - перехід до кроку 2, інакше (після закінчення перегляду) - перехід до кроку 4.

Крок 2. Нехай f - яскравість отриманого пікселя. В якості базового регіону $G_{(i)}$ ($i = 0$) приймаємо об'єднання цього пікселя з тими 8-зв'язковими з ним пікселями, яскравість яких не нижче f . Далі виконується адаптивна ітераційна процедура нарощування регіону.

Крок $G0$. Приймаємо значення f в якості початкового порогу по яскравості $B_{(i)}$ для регіону $G_{(i)}$.

Крок $G1$. Вважаємо $i = i + 1$.

Крок $G2$. Знижуємо поріг $B_{(i-1)}$ і нарощуємо безліч $G_{(i-1)}$ за рахунок 8-зв'язкових пікселів з яскравістю не меншою $B_{(i)}$.

Крок $G3$. Якщо діаметр $G_{(i)}$ перевищує $diam^*$, перехід до кроку $G4$; інакше - до кроку $G1$.

Крок $G4$. Видаляємо з системи $\{G_{(i)}\}$ регіони, діаметр яких менше $diam^*$.

Приклад рівномірного побудови порога має вигляд $B_{(i)} = B_{(i-1)} - k$, де k - ціле позитивне число, що визначає швидкість нарощування регіону; з метою підвищення ефективності сегментації правило побудови порога слід формувати з урахуванням особливостей розглянутих зображень.

Крок 3. Для регіонів $\{G_{(i)}\}$ будуємо регіональну маску (по-суті – що огинає) і оцінюємо їх контрастність згідно (3). застосовуючи критерій Kr_2 до регіонів $\{G_{(i)}\}$, визначаємо, який з них визначає шукане МІ.

Якщо кілька регіонів відповідає критерію Kr_2 , вибираємо з них регіон з максимальною контрастністю.

Крок 4. Кінець.

1.2 Виділення рухомих об'єктів у відеопотоці

Окремо слід виділити групу методів, орієнтовану на виявлення рухомих об'єктів в відеопотоці. Саме ознака руху об'єкта дозволяє виділити його на кадрі з відеопотоку і при необхідності виконати його супровід. Для оцінки ознаки використовується поле векторів руху, одержуваних методом поєднання блоків, на які розбиваються окремі кадри в стандартах відеокодування MPEG1-2 і H261 / 262/263.

Як критерій збігу блоків або цільової функції найбільш часто використовується середня абсолютна різниця яркостей пікселів порівнюваних блоків. Відома велика кількість алгоритмів, що дозволяють знаходити відповідні блоки в сусідніх кадрах відеопотоку без повного перебору шляхом визначення максимуму цільової функції. Це алгоритми, засновані на припущенні про унімодальне цільової функції, наприклад, пошук по квадрантам, алгоритми, що враховують можливість повільного руху, наприклад, трешаговий алгоритм, ієрархічні алгоритми.

Поява аномальних векторів при відсутності руху в блоці пов'язано з наявністю шуму, що викликає зміна цільової функції. Для їх компенсації проводиться поділ блоків кадру на фонові, тобто нерухомі, і рухомі, між

якими далі і знаходиться відповідність. До фоновим блокам відносяться ті, у яких рівень кореляції з відповідним блоком сусіднього кадру більше деякого порога. Після компенсації фонових блоків виконується сегментація рухомих блоків за випадковим збігом напрямки їх руху.

1.3 Згортальні нейронні мережі

Згортальна нейронна мережа (англ. Convolutional neural network, CNN) - спеціальна архітектура штучних нейронних мереж, запропонована Яном Лекуном в 1988 році і націлена на ефективне розпізнавання образів, входить до складу технологій глибокого навчання (англ. Deep learning) . Використовує деякі особливості зорової кори, в якій були відкриті так звані прості клітини, що реагують на прямі лінії під різними кутами, і складні клітини, реакція яких пов'язана з активацією певного набору простих клітин. Таким чином, ідея свёрточних нейронних мереж полягає в чергуванні свёрточних шарів (англ. Convolution layers) і субдискретизуючих шарів (англ. Subsampling layers або англ. Pooling layers, шарів підвибірки). Структура мережі - односпрямована (без зворотних зв'язків), принципово багат шарова. Для навчання використовуються стандартні методи, найчастіше метод зворотного поширення помилки. Функція активації нейронів (передавальна функція) - будь-яка, за вибором дослідника.

Назва архітектура мережі отримала через наявність операції згортки, суть якої в тому, що кожен фрагмент зображення множиться на матрицю (ядро) згортки поелементно, а результат підсумовується і записується в аналогічну позицію вихідного зображення.

1.3.1 Архітектура і принцип роботи згортальної нейронної мережі

У звичайному перцептроні, який представляє собою повнозв'язну нейронну мережу, кожен нейрон пов'язаний з усіма нейронами попереднього

шару, причому кожна зв'язок має свій персональний ваговий коефіцієнт. У свёрточній нейронній мережі в операції згортки використовується лише обмежена матриця ваг невеликого розміру, яку «рухають» по всьому оброблюваному шару (на самому початку - безпосередньо по вхідному зображенню), формуючи після кожного зсуву сигнал активації для нейрона наступного шару з аналогічною позицією. Тобто для різних нейронів вихідного шару використовуються одна і та ж матриця ваг, яку також називають ядром згортки. Її інтерпретують як графічне кодування якої-небудь ознаки, наприклад, наявності похилої лінії під певним кутом. Тоді наступний шар, що вийшов в результаті операції згортки такою матрицею ваг, показує наявність даної ознаки в оброблюваному шарі і її координати, формуючи так звану карту ознак (англ. Feature map). При цьому ядра згортки не закладаються дослідником заздалегідь, а формуються самостійно шляхом навчання мережі класичним методом зворотного поширення помилки. Прохід кожним набором ваг формує свій власний примірник карти ознак, роблячи нейронну мережу багатоканальною (багато незалежних карт ознак на одному шарі). Також слід зазначити, що при переборі шару матрицею ваг її пересувають зазвичай не на повний крок (розмір цієї матриці), а на невелику відстань. Так, наприклад, при розмірності матриці ваг 5×5 її зрушують на один або два нейрона (пікселя) замість п'яти, щоб не «переступити» шуканий ознака (рис. 1.16).

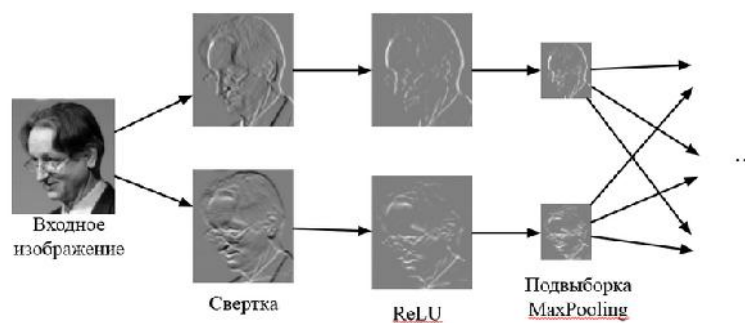


Рисунок 1.16 – Згортальна нейронна мережа

Операція Субдискретизація (англ. Subsampling, англ. Pooling, також перекладається як «операція підвибірки» або операція об'єднання), виконує зменшення розмірності сформованих карт ознак. У даній архітектурі мережі вважається, що інформація про факт наявності шуканого ознаки важливіше точного знання його координат, тому з кількох сусідніх нейронів карти ознак вибирається максимальний і приймається за один нейрон ущільненої карти ознак меншої розмірності. За рахунок цієї операції, крім прискорення подальших обчислень, мережа стає більш інваріантною до масштабу вхідного зображення.

Розглянемо типову структуру згортальної нейронної мережі більш детально. Мережа складається з великої кількості шарів. Після початкового шару (вхідного зображення) сигнал проходить серію свёрточних шарів, в яких чергується власне згортка і Субдискретизація (пулінг). Чергування шарів дозволяє складати «карти ознак» з карт ознак, на кожному наступному шарі карта зменшується в розмірі, але збільшується кількість каналів. На практиці це означає здатність розпізнавання складних ієрархій ознак. Зазвичай після проходження декількох шарів карта ознак вироджується в вектор або навіть скаляр, але таких карт ознак стають сотні. На виході свёрточних шарів мережі додатково встановлюють кілька шарів повної нейронної мережі (перцептрон), на вхід якого подаються кінцеві карти ознак (рис. 1.17).

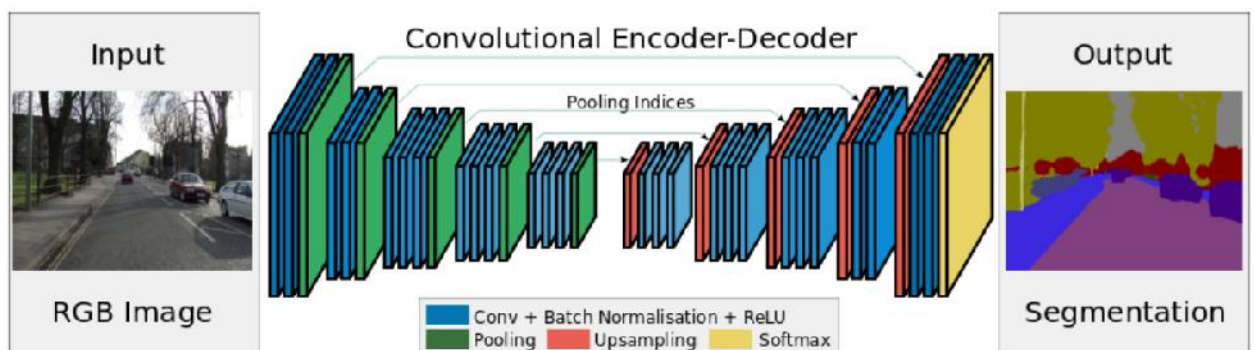


Рисунок 1.17 – Структура згорткової нейронної мережі

1.3.2 Робота згортальної нейронної мережі на прикладі сегментації сітківки ока за допомоги нейромережі (U-net)

Розглянемо реалізацію згортальної нейронної мережі, яка використовується для сегментації судин у зображеннях очного дна сітківки. Це завдання бінарної класифікації: нейронна мережа прогнозує, чи кожен піксель на зображенні очного дна є судиною чи ні.

Продуктивність цієї нейронної мережі перевіряється на базі даних DRIVE, і вона досягає найкращого балу за площею під кривою ROC порівняно з іншими опублікованими досі методами. Також у наборах даних STARE цей метод повідомляє про одне з найкращих показників (рис. 1.18).

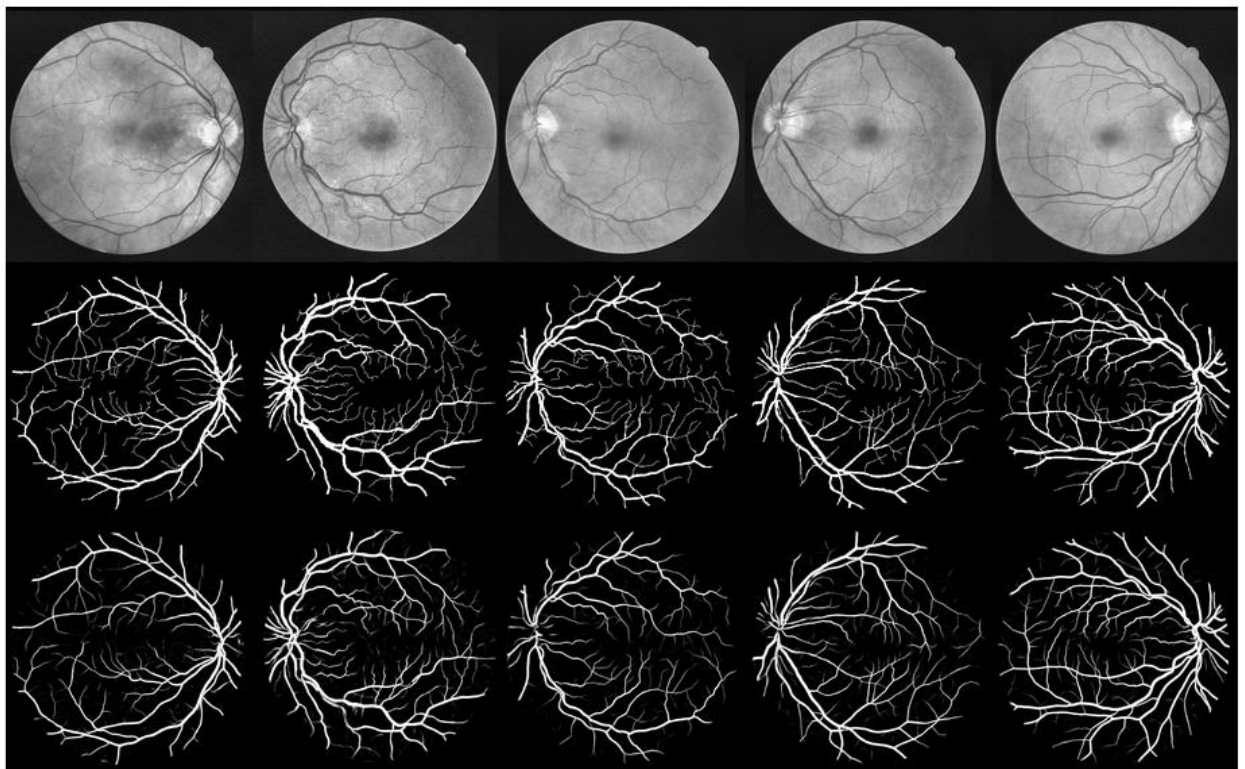


Рисунок 1.18 – Зображення сітківки ока та його аналіз після сегментації

Перед навчанням 20 зображень наборів даних DRIVE попередньо обробляються з наступними перетвореннями:

- перетворення сірого масштабу;
- Стандартизація;
- контрастне адаптивне вирівнювання гістограми з обмеженим контрастом (CLANE);
- налаштування гамми.

Тренування нейронної мережі виконується на підобразах (патчах) попередньо оброблених повних зображень. Кожен патч розміром 48x48 отримують шляхом випадкового вибору його центру всередині повного зображення. Також вибираються патчі, які частково або повністю знаходяться поза полем зору (FOV), таким чином нейронна мережа вчиться розрізняти кордон FOV від судин.

Набір 190000 патчів виходить шляхом випадкового вилучення 9500 патчів у кожному з 20 навчальних зображень DRIVE. Хоча патчі перекриваються, тобто різні патчі можуть містити однакову частину вихідних зображень, подальше збільшення даних не проводиться. Перші 90% набору даних використовуються для навчання (171000 патчів), а останні 10% - для перевірки (19000 патчів).

Нейромережева архітектура походить від архітектури U-net (див. Статтю). Функція втрат - це перехресна ентропія, а для оптимізації використовується стохастичний градієнтний спуск. Функцією активації після кожного згорткового шару є випрямний лінійний блок (ReLU), а випадання 0,2 використовується між двома послідовними згортковими шарами.

Навчання проводиться протягом 150 епох, розмір міні-партії - 32 патчі. За допомогою GeForce GTX TITAN GPU навчання триває близько 20 годин.

Тестування проводиться за допомогою 20 зображень набору даних тестування DRIVE, використовуючи золотий стандарт як основну істину. Розглядаються лише пікселі, що належать до FOV. FOV ототожнюється з масками, включеними в базу даних DRIVE.

Для підвищення продуктивності, ймовірність судна кожного пікселя отримується шляхом усереднення декількох прогнозів. З кроком у 5 пікселів

як по висоті, так і по ширині, кілька послідовних латок, що перекриваються, витягуються на кожному тестовому зображенні. Потім для кожного пікселя ймовірність судини отримується шляхом усереднення ймовірностей.

Результати, повідомлені в папці `./test`, відносяться до навченої моделі, яка повідомила про мінімальні втрати перевірки. Папка `./test` включає:

Модель:

- `test_model.png` (схематичне зображення нейронної мережі);
- `test_architecture.json` (опис моделі у форматі json);
- `test_best_weights.h5` (моделі, яка повідомила про мінімальні втрати перевірки, як файл HDF5);
- `test_last_weights.h5` (моделі останньої епохи (150-та), як файл HDF5);
- `test_configuration.txt` (конфігурація параметрів експерименту).

Результати експерименту:

- `performance.txt` (підсумок результатів тесту, включаючи матрицю плутанини);
- `Precision_recall.png` (графік відкликання точності та відповідна площа під кривою (ППК));
- `ROC.png` (крива експлуатаційної характеристики приймача (ЕХП) та відповідна ППК);
- `all _ *.png` (20 зображень попередньо оброблених оригіналів, основної правди та прогнозів відносно набору даних тестування DRIVE);
- `sample_input _ *.png` (зразок із 40 патчів попередньо оброблених оригінальних навчальних зображень та відповідної основної істини);
- `test_Original_GroundTruth_Prediction *.png` (зверху вниз, оригінальне попередньо оброблене зображення, основну правду та прогноз. У передбачуваному зображенні кожен піксель показує передбачувану ймовірність судини, поріг не застосовується).

У наступній таблиці порівнюється цей метод з іншими останніми методами, які опублікували свої показники щодо площі під кривою ROC (ППК ЕХП) на наборі даних DRIVE (рис. 1.19).

Method	AUC ROC on DRIVE
Soares et al [1]	.9614
Azzopardi et al. [2]	.9614
Osareh et al [3]	.9650
Roychowdhury et al. [4]	.9670
Fraz et al. [5]	.9747
Qiaoliang et al. [6]	.9738
Melinscak et al. [7]	.9749
Liskowski et al.^ [8]	.9790
this method	.9790

Рисунок 1.19 – Пояснення до методу

1.3.3 Оцінка навченої моделі

Продуктивність навченої моделі оцінюється на основі даних тестування drive, що складається з 20 зображень (стільки, скільки в навчальному наборі). Параметри тестування можна знову настроїти у файлі configuration.txt, зокрема в розділі [testing settings], як описано нижче:

- best_last: вибрати модель для прогнозування на тестовому наборі даних: best = модель з найнижчою втратою валідації, отриману під час тренінгу; last = модель в останню епоху.
- full_images_to_test: кількість повних зображень для тестування, не більше 20.
- N_group_visual: виберіть кількість зображень на рядок у збережених фігурах.
- average_mode: якщо є істинним, прогнозована ймовірність судна для

кожного пікселя обчислюється шляхом усереднення прогнозованої ймовірності для декількох патчів, що перекриваються.

- `stride_height`: релевантно лише якщо `true_mode` - `True`. Штрих по висоті для накладок, що перекриваються, менший крок дає більшу кількість патчів.

- `stride_width`: те саме, що і `stride_height`.

- `pothup`: стандартний висновок під час прогнозування перенаправляється та зберігається у файлі журналу.

Розділ `[experiment name]` повинен бути назвою експерименту, який ви хочете протестувати, тоді як `[data paths]` містять шляхи до тестових наборів даних. Тепер розділ `[training settings]` буде проігноровано.

Виконати тестування: `python run_testing.py`. Якщо є, буде використаний GPU. Наступні файли будуть збережені в папці з однойменною назвою експерименту:

- The ROC curve (png);
- The Precision-recall curve (png);
- Picture of all the testing pre-processed images (png);
- Picture of all the corresponding segmentation ground truth (png);
- Picture of all the corresponding segmentation predictions (png);
- One or more pictures including (top to bottom): original pre-processed image, ground truth, prediction;
- Report on the performance.

Усі результати посилаються лише на пікселі, що належать до FOV, вибрані за допомогою масок, включених у базу даних DRIVE.

Ця нейронна мережа була протестована також на іншій загальній базі даних, STARE. Нейронна мережа ідентична, як і в експерименті з набором даних DRIVE, проте деякі зміни в коді та методології були необхідні через відмінності між двома наборами даних.

STARE складається з 20 зображень очного дна сітківки з двома наборами ручної сегментації, наданими двома різними спостерігачами,

причому перший розглядався як основна істина. Навпаки на набір даних DRIVE, не існує стандартного поділу на зображення поїздів і тестів, тому експеримент проводився методом "відхід один". Цикл тренувального тестування повторювався 20 разів: при кожній ітерації одне зображення залишалося з навчального набору, а потім використовується для тесту.

Попередня обробка застосовується для набору даних DRIVE, і 9500 випадкових патчів розміром 48x48 пікселів витягуються з кожного з 19 зображень, що утворюють навчальний набір. Також область поза FOV була розглянута для вилучення пластиру. З цих патчів 90% (162450 патчів) використовуються для тренувань, а 10% (18050 патчів) використовуються для перевірки. Параметри тренувань (епохи, розмір партії ...) такі ж, як у експерименті DRIVE.

Тест виконується кожного разу на одному зображенні, що залишається поза навчальним набором даних. Аналогічно набору даних DRIVE, ймовірність судна кожного пікселя отримується шляхом усереднення по декількох перекриваючих патчах, отриманих з кроком 5 пікселів як по ширині, так і по висоті. Розглядаються лише пікселі, що належать до FOV. Цього разу ідентифікується FOV, застосовуючи поріг кольору в оригінальних зображеннях, оскільки в наборі даних STARE немає масок.

2 МЕТОДИ ВИРІШЕННЯ ЗАДАЧІ СЕГМЕНТАЦІЇ МАЛОРОЗМІРНИХ ОБ'ЄКТІВ У РІЗНИХ УМОВАХ

2.1 Метод автоматичного виявлення та вистежування багатьох малорозмірних об'єктів в умовах в умовах апіорної невизначеності.

Основним завданням систем спостереження та управління є виявлення малорозмірних і точкових динамічних об'єктів і оцінка їх траєкторії (простежування) в часі. В роботі розглянуті питання виявлення малорозмірних об'єктів по виявленим треках при низькому відношенні сигнал/шум на підставі стробовий методу. Цей метод можна віднести до так званих методів "супроводу до виявлення" (Track Before Detect, TBD). На відміну від класичного супроводу, при якому рішення про наявність об'єкта приймається по кожному кадру на основі внутрікадрової обробки, при використанні методу TBD аналізується обмежена послідовність кадрів, і рішення приймається на підставі накопиченої інформації за останнім кадром послідовності. Обробка проводиться по ковзної послідовності кадрів, результат обробки формується по кожному кадру (з початковою затримкою, яка визначається довжиною послідовності).

Первинна внутрішньокadroва обробка призначена для виявлення точкових об'єктів на основі автоматичного виділення областей целеподобних об'єктів, адаптивного порогового обмеження і оцінки центра ваги. В результаті первинної обробки формується бінарне зображення, в якому координатам малорозмірних об'єктів або перешкод відповідає значення 1, а фону - 0. Це зображення поряд з об'єктами Містить значну кількість перешкод. Далі виконується виявлення нових траєкторій руху, підтвердження існуючих або завершення/видалення траєкторій на підставі даних, накопичених за обмеженої послідовності кадрів. На підставі аналізу приймається рішення про наявність об'єктів інтересу (рис. 2.1).

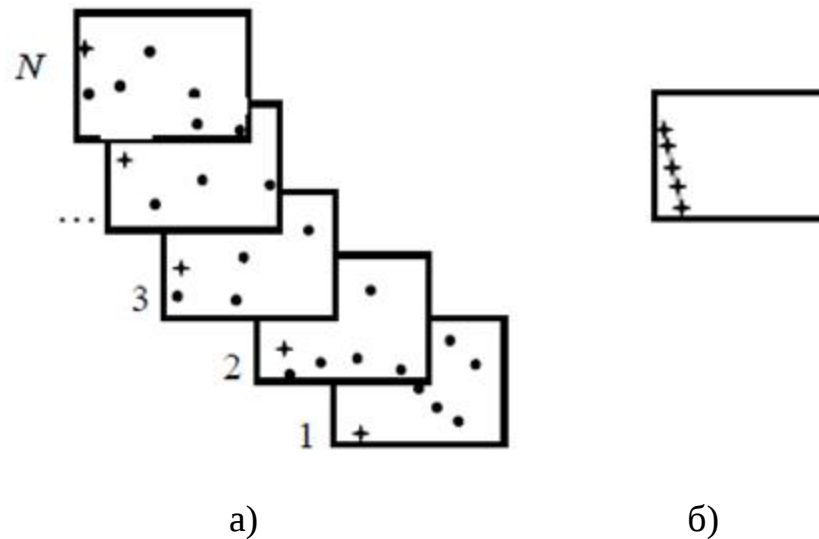


Рисунок 2.1 – Приклад формування треку по послідовності з N кадрів (а - послідовність кадрів, б - накопичений трек; хрестиком відзначені об'єкти, точками - перешкоди)

При використанні відеосистем кругового огляду динамічний діапазон міжкадрових зсувів малорозмірних об'єктів істотно збільшується в порівнянні з системами секторного спостереження. Розширення діапазону можливих швидкостей спостережуваних об'єктів визначає збільшення розміру стрибкає. В цьому випадку в протиріччя входять вимоги збільшення розміру стрибкає і забезпечення необхідної ймовірності правильного виявлення при заданій ймовірності помилкової тривоги. Дійсно, при збільшенні стрибкає зростає число відміток, які формують помилкові треки, тобто підвищується ймовірність помилкової тривоги. Для забезпечення заданого рівня помилкових тривог необхідно зменшити число відміток в кадрі, що досягається винятком відміток при первинній обробці з меншим відношенням сигнал / шум, це призводить до пропуску об'єктів з малою інтенсивністю сигналу. Для можливості роботи у великому динамічному діапазоні швидкостей об'єктів авторами розроблений метод, який є розвитком стробовий методу.

2.1.1 Опис роботи методу

Вхідними даними є послідовність бінарних зображень, точкові відліки яких можуть належати або об'єкту, або заваді. Координати одиничних відліків визначають безліч вимірів

$$Z(k) = \{Z_j^k\}_{j=1}^{M_k}, \forall k = [1, N], \quad (2.1)$$

де k - поточний номер кадру, N - число кадрів в ковзній послідовності, M_k - число одиничних відміток в k -м кадрі. Завдання полягає в тому, щоб з множини $\{Z_j^k\}$ вибрати ті відліки, які формують трек.

Рішення про виявлення об'єкта приймається за відсутності апіорних даних про час появи і зникнення, кількості, положенні і параметри руху (напрямок, швидкість) об'єкта при заданій максимальній швидкості руху. Рішення з оцінкою координат об'єкта має вироблятися з кожного кадру послідовності і забезпечувати мінімально можливу тимчасову затримку t до прийняття рішення

$$t = Nt_f, \quad (2.2)$$

де t_f – довжина кадра.

Вектор стану об'єкта x_k визначає координати об'єкта і оцінку вектора швидкості по кадру k в момент часу t_k . Передбачається, що швидкість і напрямок руху об'єкта не змінюються миттєво в часі, проводиться кусочно-лінійна апроксимація траєкторії руху об'єкта на змінному часовому інтервалі t . На рис. 2.2 представлений фрагмент зображення з накопиченої в часі траєкторією руху об'єкта. Точками позначені оцінки координат (y, x) в послідовності кадрів $[1, \dots, N]$ тривалістю t (на рис. 2.2, $N = 4$). При кусково-лінійної апроксимації рух об'єкта від кадру до кадру відбувається з

постійною швидкістю і описується лінійним рівнянням

$$x_{k+1} = F_k x_k, \quad (2.3)$$

Де $x_k = [r_k \ c_k \ v_{rk} \ v_{ck}]'$; r_k – номер строки, c_k – номер стовпця; v_{rk} , v_{ck} – швидкість руху об'єкта по рядку і стовпцю відповідно; «'» - знак транспонування; F_k – функція, що описує траєкторію (матриця переходу):

$$F_k = \begin{bmatrix} 1 & 0 & t_f & 0 \\ 0 & 1 & 0 & t_f \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Під траєкторією розуміється безліч відліків, отриманих в N послідовних кадрах і задовольняють рівняння стану системи (2.3)

$$T_i^k = \{z_{i_1}^{k-N+1}, z_{i_2}^{k-N+2}, \dots, z_{i_{n-1}}^{k-1}, z_{i_n}^k\}, \quad (2.4)$$

де i – номер траєкторії $[i_1, \dots, i_n]$ – номер виміру, $i_n \leq N$.

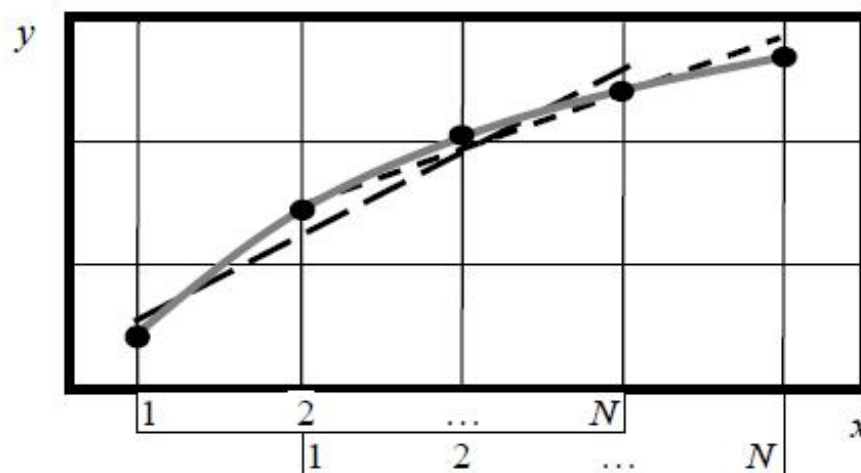


Рисунок 2.2 – Точки у системі координат у послідовності кадрів

Завдання визначення приналежності об'єкта траєкторії вирішується по кожному кадру в реальному часі на основі аналізу послідовності довжиною N кадрів. Підтвердження траєкторії на N кадри дозволяє прийняти рішення про те, що відлік відноситься до об'єкту, а не заваді. Належність траєкторії оцінюється при допустимому відхиленні оцінки вектора стану в кадри k від очікуваного треку, обумовленому "не жорстким" растром в разі кругового огляду.

2.1.2 Стробо́вий метод виявлення малорозмірних об'єктів

Цей метод побудований на аналізі зображення в двох строб. (Узагальнена структурна схема представлена на рис. 2.3).

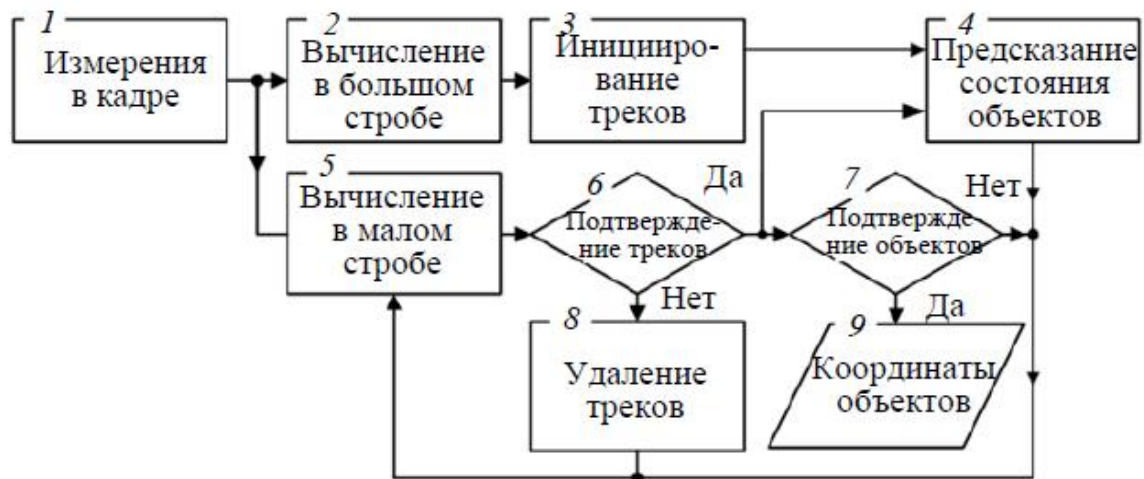


Рисунок 2.3 – Узагальнена структурна схема

Розмір першого (великого) строб визначається діапазоном можливих швидкостей об'єктів, алгоритм виконується так само, як в роботі. Цей стрибак призначений для ініціалізації треку малорозмірних (точкових) об'єктів. У даній роботі під ініціалізацією треку розуміється визначення по двох суміжних кадрах координат можливих об'єктів інтересу, що задовольняють динамічному діапазону швидкостей. В якості запобіжного

відстані між об'єктами використовується евклидова міра. Необхідно підтвердити наявність треку на послідовності ще $N-2$ кадрів. Розмір великого строба визначається тим, що положення рухомого об'єкту в поточному кадрі з точністю до швидкості його руху не виходить за межі строба, центр якого сполучений з координатами об'єкта в попередньому кадрі. Всі відліки, що належать стробу, ініціюють новий трек. Оскільки час появи об'єкта в кадрі апіорі не відомо, кожна позначка може виявитися об'єктом, тому положення центру великого строба визначається координатами всіх відміток в попередньому кадрі (блоки 2, 3).

Координати центру другого (малого) строба оцінюються за передбаченням положення об'єкта в k -м кадрі на підставі інформації, отриманої в $(k-1)$ -м кадрі (див. (2)) для всіх активних (ініційованих і підтверджених) треків (блок 4). Розмір малого строба визначається можливою нестабільністю раstra, помилками оцінки координат і допустимим прискоренням об'єкта. У другому Строба (блок 5) виконується аналіз з метою підтвердження треків, ініціалізованих в попередньому кадрі і простежується, завершення/видалення непідтверджених треків. Підтверджується той трек, для якого в поточному кадрі в малому Строба знаходиться відмітка (блок 6). Для мінімізації ймовірності пропуску цілі невизначеність, що виникає при наявності декількох відміток в малому Строба, дозволяється розгалуженням треку і спадкуванню передісторії треку. Обробка в малому Строба дозволяє істотно зменшити число ініційованих треків. Додаткове зниження ймовірності помилкових тривог досягається перевіркою лінійності треку, необхідність цієї перевірки обумовлена розміром малого строба. Для цього за методом найменших квадратів по N відліку визначаються параметри прямої лінії і оцінюється відхилення відліків від лінії для кожної з N координат. Оцінки стану треків оновлюються для кожного кадру.

Такий підхід дозволяє аналізувати рухомі об'єкти (при великому діапазоні швидкостей) без зростання ймовірності помилкової тривоги,

неминуче виникає внаслідок суттєвого збільшення розміру стрибкає при роботі за методом. Число ініційованих треків повинно зростати при збільшенні числа помилкових відміток, однак введення малого стрибкає дозволяє послабити цю залежність. У табл. 2.1 приведена залежність сумарного числа помилкових тривог за весь час спостереження від розміру великого стрибкає для послідовності з 50 кадрів (1024×512 елементів), при наявності 1 об'єкту інтересу і 100 перешкод в кожному кадрі.

Таблиця 2.1 – Залежність числа помилкових цілей від розміру великого строба

Розмір великого строба, елементи	41×41	61×61	81×81	101×101	121×121	161×161
Число помилкових тривог	1	3	3	3	3	3

Традиційно для підтвердження треку перевіряється наявність об'єкта в M кадрах з N спостережуваних. Необхідність мінімізації часу виявлення об'єкта при заданій ймовірності помилкової тривоги тягне за собою вимогу наявності об'єкта, що належить треку, в кожному з N кадрів (блок 7), тобто при $i_n = N$ (вираз (2.3)), координати об'єкта передаються на вихід (блок 9).

Рішення про завершення/видаленні треку зазвичай приймається за відсутності об'єкта, що належить треку, протягом L кадрів підряд, часто $L = [3,4]$. В даному випадку $L = 1$. Видалення треку виробляється в блоці 8. Зрозуміло, що збільшення значення N призводить до зменшення ймовірності помилкової тривоги при заданій щільності перешкод, малі значення N навіть при некорельованій заваді породжують помилкові треки. Ця залежність проілюстрована в таблиці 2.2. Дані наведені для кадру розміром 1080×1920 елементів. Розмір великого стрибкає 161×161 , малого – 23×23 елемента. Генерація точкових відміток перешкоди виконана для випадку рівномірного

розподілу перешкоди по кадру. Загальна довжина послідовності $(1000 + N - 1)$ кадрів. Об'єкти в кадрах відсутні. З таблиці видно, наприклад, що для 300 перешкод в кожному кадрі (ймовірність помилкової тривоги $1,44 \times 10^{-4}$) при $N = 3$ були виявлені ≈ 32 неправдиві позначки на кадр, при $N = 4 - 0,987$, а при $N = 5$ виникає по одній помилковій позначці в 27 кадрах з тисячі.

Підвищення щільності перешкод при збереженні розмірів великого і малого стробов призводить до підвищення ймовірності помилкової тривоги, що показано в табл. 2.2.

Таблиця 2.2 Залежність числа помилкових цілей від довжини послідовності N при заданій інтенсивності перешкоди

Число перешкод	Число помилкових тривог на кадр		
	$N=3$	$N=4$	$N=5$
100	1,297	0,010	0
120	2,086	0,024	0
140	3,274	0,044	0
160	5,031	0,090	0
180	7,138	0,146	0,003
200	9,809	0,193	0,004
220	12,768	0,305	0,008
240	16,435	0,411	0,018
260	20,699	0,610	0,016
280	25,660	0,788	0,024
300	31,550	0,987	0,027

2.1.3 Дослідження ефективності методу

Були синтезовані відеопослідовності треків при задається числі відміток перешкод для сценаріїв з одним, двома або багатьма об'єктами, для випадку перетину траєкторій руху об'єктів і розщеплення траєкторії.

Траєкторія задавалася лінійною функцією при розподіленій за нормальним законом помилку координат. Координати перешкодних відміток розподілялися в кадрі по рівномірному закону. При синтезі послідовності формувалася тестовий список координат об'єктів для кожного кадру. Розмір послідовності становив 25-100 кадрів. Оцінка ефективності проводилася шляхом зіставлення тестового списку та списку виявлених за розробленим методом об'єктів. Крім цього, формувалося накопичене протягом послідовності зображення отриманих треків. На рис. 4 представлено накопичене зображення координат виявлених об'єктів для семи целеподобних відміток, дві з яких (1 і 7) нерухомі, лише схильні до випадковим зміщення в допуски помилок оцінок координат об'єкту, нестабільності растра та ін. (рис. 2.4).

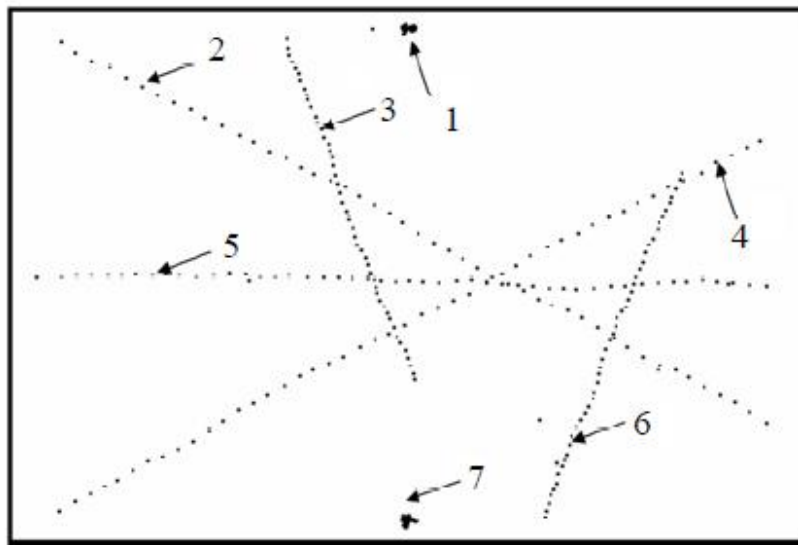


Рисунок 2.4 – Накопичене зображення координат виявлених об'єктів

Розмір тестового зображення 768×512 , синтезовано 75 помехових відміток в кожному кадрі послідовності. Довжина послідовності 50 кадрів, довжина послідовності для прийняття рішення - 4 кадри. В експерименті треки всіх об'єктів простежено.

2.1.4 Підсумок за методом автоматичного виявлення та вистежування багатьох малорозмірних об'єктів в умовах в умовах апріорної невизначеності

Метод виявлення об'єктів, заснований на виявленні треку з використанням стробовий методу. У методі використано припущення, що на невеликій послідовності відеокадрів трек можна апроксимувати лінійною функцією. Виявлення об'єкту проводиться після аналізу N кадрів послідовності. Час виникнення і зникнення об'єкта невідомо, число виявлених об'єктів не обмежена і невідомо. Метод ефективний при великому динамічному діапазоні можливих швидкостей об'єктів інтересу. Скорочення тривалості послідовності знижує вимоги до лінійного характеру траєкторії, оскільки лінійність повинна виконуватися лише в ковзному вікні короткої послідовності кадрів.

2.2 Ідентифікаційний метод розпізнавання образів малорозмірних підводних об'єктів

Описано метод розпізнавання образів підводних об'єктів, заснований на використанні ідентифікаційних вимірів, в результаті яких виходить число, кількісно характеризує форму будь-якої характеристики інформативного сигналу. Імовірність розпізнавання і класифікації підводних об'єктів визначається повнотою заздалегідь створеної бази образів відомих об'єктів, сформованої відповідно до оцінок сертифікованих експертів. При базі даних в три образи ймовірність розпізнавання склала більше 80%. Отримані результати, висока чутливість і відсутність помітних крайових ефектів методу вказують на можливість реалізації для гідролокатором бічного огляду принципово нового типу програмного забезпечення, яке дозволить підвищити їх інтелектуалізацію в області розпізнавання і класифікації підводних об'єктів.

Гідрологічні дослідження водних акваторій є важливою складовою вирішення завдання забезпечення безпеки судноплавства. Основною метою при цьому ставиться виявлення і розпізнавання таких об'єктів, що знаходяться у водному середовищі, які потенційно несуть небезпеку для різних судів. Прикладами таких об'єктів можуть бути затонулі судна, підводні скелі, кабельні лінії, трубопроводи, міни, різні контейнери та інші об'єкти. При проведенні гідрологічних досліджень основними інструментами є гідролокатори. Найбільш поширені гідролокатори бокового огляду (ГБЗ). Одним з перспективних напрямків у розвитку сучасного гідролокаційних обладнання є застосування в них фазованих антенних решіток (ФАР) випромінювачів і приймачів ультразвукових хвиль. Такий підхід спільно з застосуванням лінійної частотної модуляції дозволяє значно підвищити роздільну здатність і стійкість при виявленні об'єктів [5, 6].

2.2.1 Задачі що вирішує ідентифікаційний метод

В даний час в ГБЗ при обробці луна-сигналу, отриманого від об'єкта дослідження, як правило, використовується швидке перетворення Фур'є (ШПФ). Даний підхід має кілька недоліків при низькочастотному зондуванні:

- 1) на точність вимірювання впливають спектральні складові шумів виду $1 / F$, де F - частота сигналу;
- 2) для отримання необхідного дозволу необхідно збільшувати час вимірювання;
- 3) не вирішено питання автоматичної класифікації і розпізнавання об'єктів на фоні перешкод на площині, що сканується.

Підвищення ж частоти зондуючого сигналу спричиняє збільшення втрат потужності, і зменшення дальності виявлення об'єктів. Виділення об'єкта і подальше його розпізнавання в даний час здійснюється в ручному режимі експертами в цій галузі, які застосовують методи цифрової обробки зображень, вейвлет-аналіз і нейронні мережі, моделювання випадкового поля

і нечіткої логіки, адаптивні системи пошуку. Усунення перерахованих недоліків призводить до ускладнення програмного забезпечення ГБЗ, збільшення часу аналізу і часу прийняття рішення, і, як наслідок, до збільшення витрат на проведення досліджень.

Для розпізнавання образів підводних об'єктів пропонується використовувати технології ідентифікаційних вимірів, які засновані на кількісному оцінюванні форми досліджуваних об'єктів. Доведена можливість кількісного оцінювання (вимірювання) форми характеристик сигналів.

До теперішнього часу сформований істотний науковий доробок, що дозволяє говорити про те, що в теорії і техніці вимірювань з'явилося новий науковий напрям - ідентифікаційні виміри (ІІ) сигналів. Засновником цього напрямку по праву вважається доктор технічних наук, професор Ю. М. Клікушін. Суть ІІ полягає в кількісному оцінюванні форми характеристик сигналів.

Ключовою особливістю ІІ є інтегральне вимірювання форми характеристик сигналів. При цьому сам сигнал (або його характеристика) розглядається як щось ціле, неподільне, наприклад, як єдиний масив значень, якому за певними правилами встановлюється відповідно число, назване ідентифікаційним. Це число є результатом вимірів. Як показують проведені раніше дослідження, подібний підхід в більшості випадків дає відчутний наочний результат.

При застосуванні інтегральних методів слід враховувати неоднозначність: одне і теж ідентифікаційне число може характеризувати різні сигнали. Ця особливість всіх інтегральних перетворень. Наприклад, одні і ті ж числові значення точкових характеристик випадкових сигналів можуть бути отримані від різної сигналів.

Одним з можливих шляхів зменшення неоднозначності розпізнавання може служити вимір внутрішньої структури сигналів, що дозволяє розрізнити, наприклад, два сигнали, що мають однакові інтегральні ідентифікаційні показники, але різняться видом їх розподілу по всій

тривалості сигналу. Крім того, дослідження внутрішньої структури сигналів шляхом їх сканування по довжині дозволяє виявляти, ідентифікувати і класифікувати приховані закономірності у вигляді деяких нечітко визначених кластерів. В теорії II виділення подібних кластерів здійснюється на основі вимірювання розподілів миттєвих значень (РМЗ) відповідних сигналів.

Ця обставина є принциповою відмінністю теорії II від класичних підходів, що реалізуються в задачах аналізу сигналів, оскільки представляє нову концепцію в розпізнаванні образів - переходу від вимірювань моментів до вимірювання РМЗ.

Базова ідея теорії II пояснюється на рис. 2.5, де представлена технологія виконання ідентифікаційних вимірів. Вибіркова реалізація вхідного сигналу $X(t)$ довільного обсягу N надходить на вхід ідентифікаційного перетворювача (тестера), на виході якого формується ідентифікаційне число $G = Id [X(t)]$.

Це число порівнюється з багатозначною мірою - ідентифікаційної шкалою (ІШ) подібних же чисел ($G_1, G_2, \dots G_k$), отриманих в результаті попередньої калібрування. Основною властивістю ідентифікаційних чисел є їх масштабна інваріантність - незалежність від параметрів математичного очікування і дисперсії аналізованого сигналу.

В результаті раніше проведених досліджень [11, 12] була встановлена фундаментальна закономірність (табл. 2.3), що виявляється в тому, що ідентифікаційні числа (K - параметр варіабельності, A - параметр форми розподілу, F_c - квазічастота сигналу) впорядковують симетричні розподілу випадкових сигналів - двумодальное (2мод), арксінусное (АРКС), рівномірний (так само), трапецієподібно (трап), трикутне (симпо), нормальне (НОРМ), Лапласа (ЛАПЛ) і Коші (КОШІ) - в діапазоні від 4 до N .

Примітка: N - обсяг вибірки реалізації сигналу; L - кількість реалізацій сигналу; t_n - час спостереження сигналу; Rank - порядковий номер позначки на шкалі; $\langle K \rangle$, $\langle A \rangle$, $\langle F_c \rangle$ - значення параметрів K , A , F_c усереднені по L реалізацій сигналу для кожної позначки шкали.

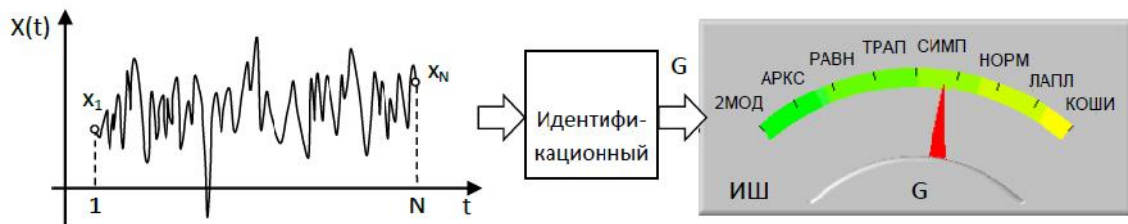


Рисунок 2.5 – Технологія виконання ідентифікаційних вимірів

Таблица 2.3 Идентификационная шкала параметров формы и variability случайных сигналов

N=10000, L=100, t _н =1с	Позначки шкалы - реперные точки (вид закона распределения случайного сигнала)							
	2МОД	АРКС	РАВН	ТРАП	СИМП	НОРМ	ЛАПЛ	КОШИ
Rank	1	2	3	4	5	6	7	8
<K>, рад	1	1,27	1,33	1,39	1,4	1,41	1,5	1,87
<A>, рад	4	6,28	8	10	12	19	36	N
<F _c >, Гц	2500	2027	1667	1350	1180	736	416	6

Таким чином, в найпростішому випадку розпізнавання невідомого вхідного сигналу полягає в інтерполяції положення стрілки показчика між оцифрованими і поименованими відмітками ІШ, із зазначенням ступеня приналежності імені досліджуваного сигналу до найближчої реперною точці.

Така інтерпретація результатів вимірювань фактично означає рішення задачі лінійної класифікації, при якій імена реперних точок позначають імена класів групування. Зокрема, для ІШ (табл.3.3), що складається з 8 реперних точок, безліч вхідних аналізованих сигналів може бути розділене по параметру форми A на 8 класів, а за двома параметрами A і F_c - на 64 класу.

Параметри ІШ можуть бути представлені у вигляді рангових характеристик в графічній і аналітичній формах (рис. 2.6, рис. 2.7). Це дозволяє проводити аналітичну інтерполяцію показань ідентифікаційного

тестера (коли стрілка покажчика знаходиться між оцифрованими відмітками реперних точок) і, тим самим, підвищувати достовірність оцінки РМЗ сигналу.

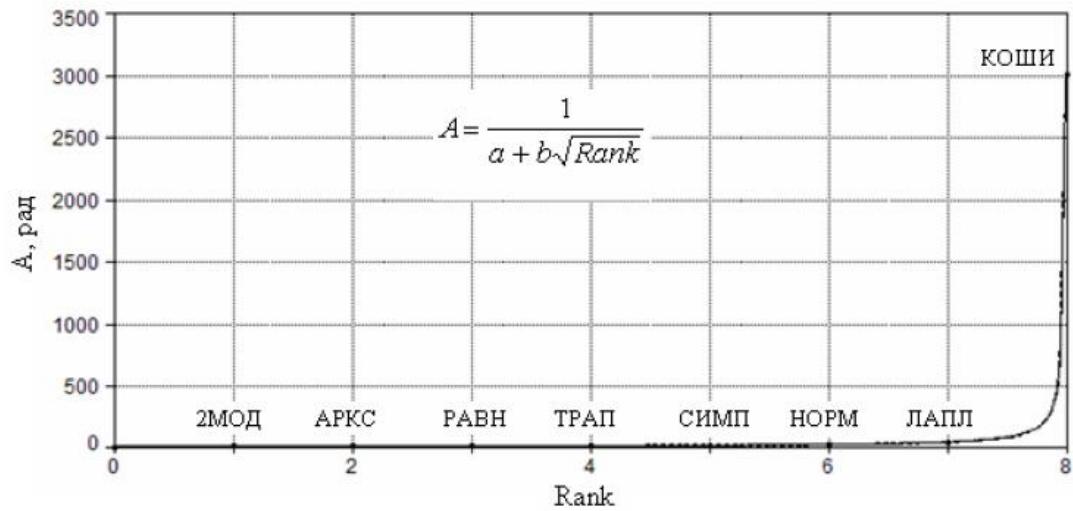


Рисунок 2.6 – Графік залежності параметра форми A від порядкових номерів $Rank$ розподілів

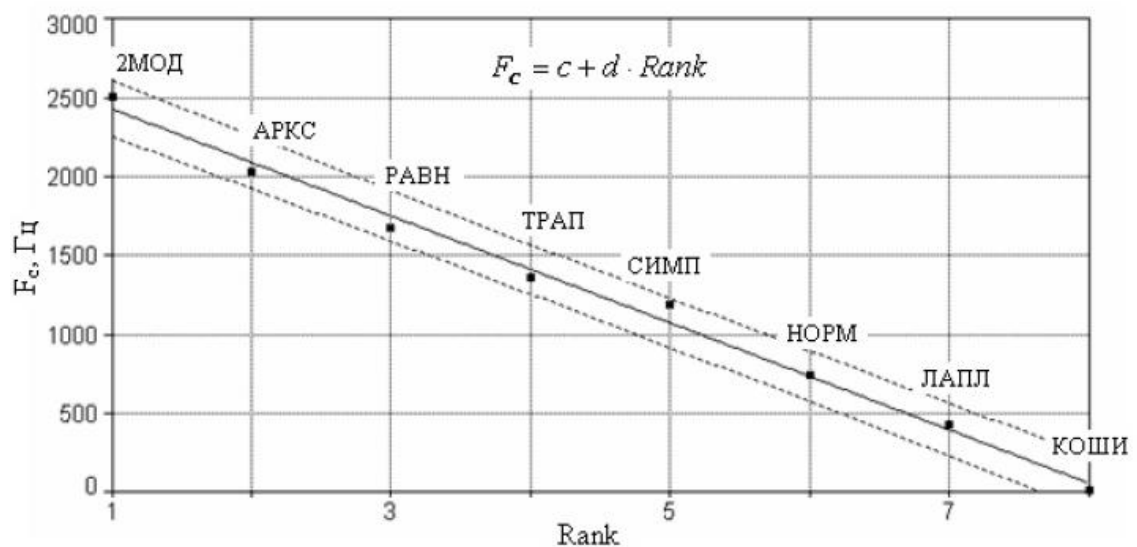


Рисунок 2.7 – Графік залежності квазічастоти F_c від порядкових номерів $Rank$ розподілів

Порівняння графіків показує, що, виходячи з критерію простоти реалізації і в силу лінійності функції перетворення, найбільш перспективним є використання квазічастоти F_c в якості ідентифікатора РМЗ випадкових сигналів. Фізичний сенс поняття квазічастоти полягає в тому, що вона характеризує середню кількість локальних однойменних екстремумів (максимумів або мінімумів) сигналу за одиницю часу спостереження. При цьому, якщо сигнал - випадковий, то квазічастота визначає вид РМЗ, якщо ж сигнал - періодичний, то квазічастота визначає значення частоти сигналу.

Якщо сигнал є деякою сумішшю регулярної та хаотичної компонент, то квазічастота повинна приймати таке значення, яке може служити мірою змісту цих компонент в сигналі. Проведені дослідження дозволили встановити нову для теорії сигналів закономірність - квазічастота F_c випадкового сигналу не менш квазічастоти F_x будь-якого іншого сигналу з тим же значенням параметра форми A , тобто таким же розподілом миттєвих значень

$$F_c \geq F_x, \text{ при } (A = A_x). \quad (2.4)$$

Використовуючи формулу (2.4), в роботі [13] було запропоновано спосіб, що дозволяє оцінювати відносний вміст регулярної компоненти складного сигналу. Для прикладу, на рисунку 2.8 показані два сигнали, один з яких (8.wav) є «чистим» (незашумленим) сигналом, а інший (7.wav) - Містить хаотичну компоненту (шум).

Результати аналізу показують, що, при однаковій формі РМЗ (ЛАПЛ | 0,85), ступінь регулярності незашумленим сигналу становить $Reg = 92\%$, а на перешкоди - $Reg = 77\%$. Таким чином, відмінність сигналів з даного ідентифікаційним показником становить 15% . Іншими словами - сигнали схожі на 85% .

Оскільки структура (рис. 2.8) Містить вбудовану базу знань у вигляді багатозначної міри РМЗ, то ідентифікаційні інструменти можуть бути

віднесені до інтелектуальних засобів вимірювання і розпізнавання сигналів. Найбільш перспективними напрямками використання ідей і методів ідентифікаційних вимірів є контроль і діагностика стану будь-яких об'єктів і процесів, представлених своїми вибірковими реалізаціями.

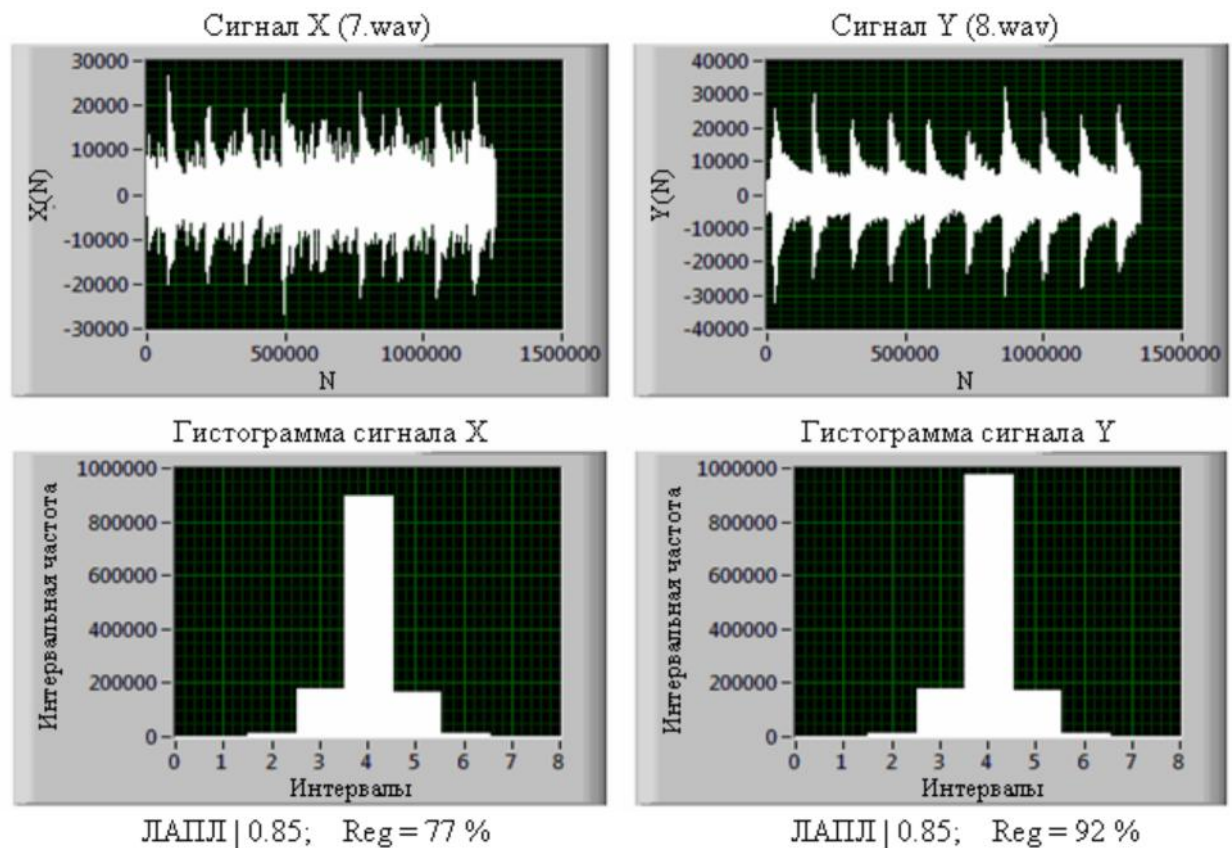


Рисунок 2.8 – Приклад порівняння двох сигналів за ступенем регулярності

2.2.2 Результати експерименту

При розпізнаванні образів підводних об'єктів задача аналізу зводиться до виявлення та локалізації в структурі сигналу деяких неоднорідностей. Тому аналізатор повинен містити скануючий пристрій, розпознаватель і пороговий фільтр. Структурна схема ідентифікаційного аналізатора (ІА) зображена на рис. 2.9. Аналізатор складається з скануючого пристрою (сканера), ідентифікаційного тестера і компаратора. Сканер послідовно

«переглядає» вхідний сигнал $X(t)$, починаючи з першого відліку. Вихідний сигнал сканера $Y(t)$ вимірюється тестером, на виході якого формується ідентифікаційне число $G(t)$.

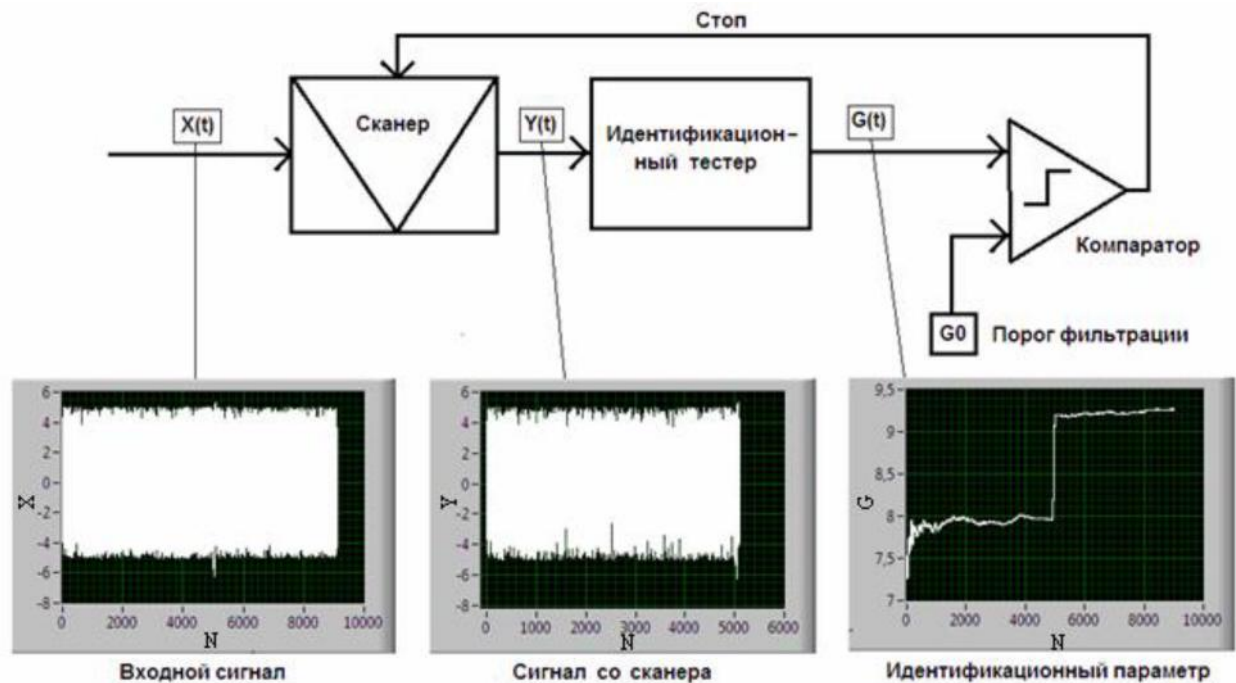


Рисунок 2.9 – Структурна схема ідентифікаційного аналізатора

Компаратор порівнює поточне значення ідентифікаційного числа $G(t)$ з деяким пороговим рівнем G_0 . При виконанні умови $G(t) \geq G_0$ фіксується момент часу $t = t_1$ спрацювання компаратора і робота сканера зупиняється. Момент часу t_1 визначає положення лівої межі неоднорідності сигналу. Аналогічно, при скануванні сигналу $X(t)$ від останнього звіту до першого можна визначити момент часу t_2 , який буде задавати положення правої межі неоднорідності сигналу. Розмір неоднорідності обчислюється, як різниця моментів часу: $t_2 - t_1$.

Оскільки розмір і положення неоднорідності по довжині сигналу вимірюються в просторі ідентифікаційних чисел $G(t)$, а не миттєвих значень $X(t)$ вхідного сигналу, то, потенційно, такий спосіб виділення області існування неоднорідності (або кластеризації) володіє більш високою

статистичною стійкістю в порівнянні з класичними амплітудними вимірами. Значення порога G_0 спрацьовування компаратора задається, виходячи з апіорних знань про РМЗ сигналів фону (середовища) і неоднорідності (об'єкта).

У процесі досліджень було проведено порівняння моделей ідентифікаційного і класичного (дисперсійного) методів розпізнавання образів об'єктів у водному середовищі.

На рис. 2.10 представлені результати такого порівняння для відносини сигнал-шум (ВСШ) рівного 1.

Вимірювання були проведені для випадку, при якому неоднорідність (модель об'єкта) представляла собою адитивну суміш випадкових сигналів з нормальним і рівномірним розподілами.

У свою чергу, ця суміш компілювати в рівномірно розподілений випадковий сигнал, який служив моделлю середовища.

Візуально подібний сигнал показаний на дисплеї «Вхідний сигнал» (рис. 5). На рис. 2.10 графіки Left-1 і Left-2 відносяться до випадку сканування положення об'єкта зліва, а графіки Right-1 і Right-2 - праворуч.

Як впливає з зіставлення образів, отриманих ідентифікаційним (дисплей Cluster-1) і класичним (дисплей Cluster-2) методами, амплітуда ідентифікаційного кластера на 13 порядків більше амплітуди кластера, отриманого шляхом стандартного дисперсійного аналізу.

У таблиці 2.4 промодельовати чисельні оцінки ймовірності правильного виявлення об'єкта при $ЗСШ = 0,03$.

За сукупністю різних розмірів (100; 250; 1000, 5000 відліків) об'єкта і його положень (на початку, середині і наприкінці) в області огляду, середнє значення ймовірності правильного виявлення становить 0,89 при кількості повторень для кожного випадку 1000.

Таким чином, ймовірність виявлення об'єкта, представленого сигналом з 1000 звітів і більш, за допомогою описаного методу склала понад 0,97.

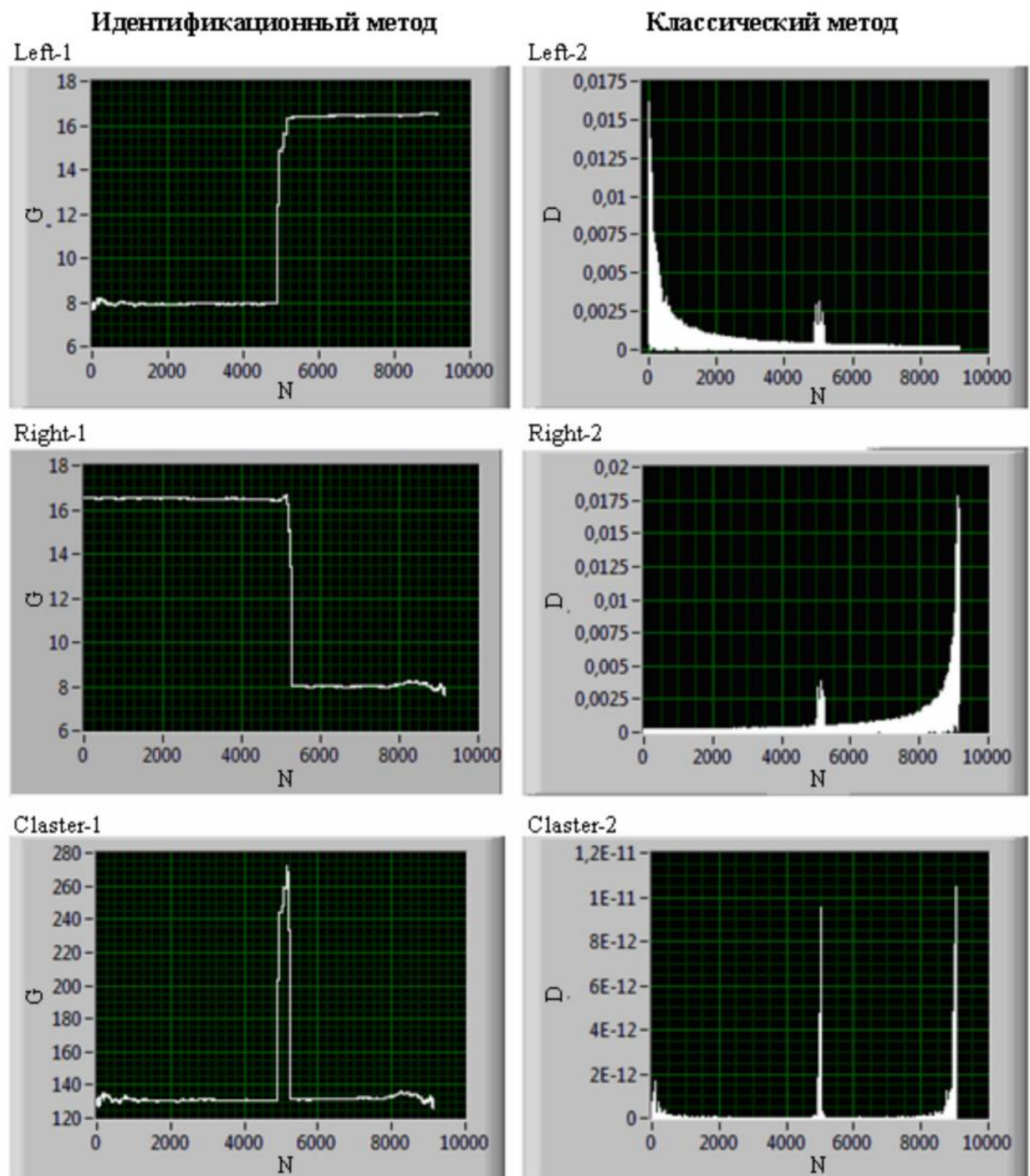


Рисунок 2.10 – Приклади виділення границь і області неоднорідності сигналів ідентифікаційним (лівий стовпець) і класичним (правий стовпець) методами для ЗСШ = 1

Описаний метод застосований для вирішення практичного завдання розпізнавання і класифікації образів підводних об'єктів, зображення яких отримані за допомогою ГБЗ. Для прикладу створення бази образів були

обрані три підводних об'єкта, названих «Судно» (З), «Торпеда» (Т), «Мина» (М). Приклади зображень таких об'єктів показані на рис. 2.11.

Таблиця 2.4 Ідентифікаційна шкала параметрів форми і варіабельності випадкових сигналів

Розмір об'єкта (В кількості відліків)	Положення об'єкта в області огляду	Імовірність правильного виявлення
100	початок	0,96
	середина	0,68
	кінець	0,4
250	початок	1
	середина	0,93
	кінець	0,85
1000	початок	0,98
	середина	0,97
	кінець	0,97
5000	початок	0,998
	середина	0,996
	кінець	0,997
Середнє значення:		0,89

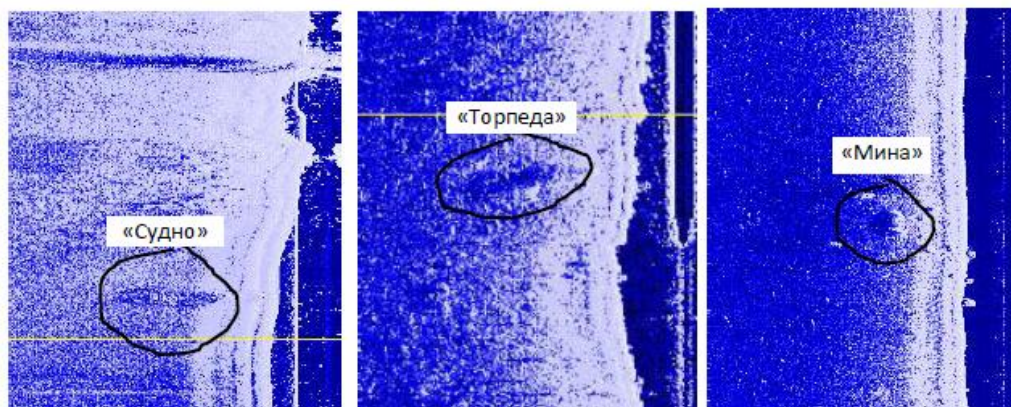


Рисунок 2.11 – Приклади образів об'єктів, отриманих за допомогою ГБЗ

В основі технології розпізнавання і класифікації образів підводних об'єктів лежить ідентифікаційна шкала параметрів форми і варіабельності випадкових сигналів (табл. 2.4). Кожен об'єкт в базі образів представлений у вигляді коду, сформованого з параметрів A , F_c . Спочатку об'єкт локалізується на тлі перешкод за допомогою алгоритму, описаного вище. Потім фрагмент зображення сканується по рядках і стовпцях. Для кожного рядка і стовпця визначаються параметри A і F_c , що формують два ідентифікаційні коди - рядковий і столбцову відповідно. Отримані коди є унікальними ідентифікаторами фрагмента зображення. В результаті порівняння ідентифікаційного коду невідомого фрагмента з базою даних (БД) формується судження про приналежність невідомого об'єкта найближчого відомому з БД.

Результати розпізнавання і класифікації образів підводних об'єктів представлені в таблиці.

Таблиця 2.5. Результати класифікації образів підводних об'єктів

№ п/п	Ім'я файла	Експертн а оцінка	Ймовірність збігу з образами БД, %					
			по коду строки			по коду стовпця		
			С	Т	М	С	Т	М
1	2	3	4	5	6	7	8	9
1	07.sff	М	16	38	65	15	39	71
2	08.sff	М -БД	10	12	87	9	15	92
3	14.sff	М	23	34	58	21	38	68
4	15.sff	М	29	25	71	16	46	57
5	17.sff	М	13	44	55	31	57	49
6	36.sff	М	31	51	45	22	45	53
7	53.sff	М	14	28	63	13	37	58
8	25.sff	М	12	33	65	20	32	70

Продовження 1 Таблиці 2.5. Результати класифікації образів підводних об'єктів

№ п/п	Ім'я файла	Експертна оцінка	Ймовірність збігу з образами БД, %					
			по коду строки			по коду стовпця		
			С	Т	М	С	Т	М
9	98.sff	М	5	46	53	28	27	51
10	99.sff	М	11	47	42	34	59	20
11	56.sff	М	23	29	58	23	35	55
12	60.sff	М	18	21	74	12	23	69
13	19.sff	М	17	36	66	29	27	65
14	34.sff	М	9	40	47	14	37	53
15	61.sff	М	34	36	67	18	22	75
16	63.sff	М	24	37	60	29	41	52
17	77.sff	М	41	36	55	23	40	46
18	93.sff	М	45	58	32	40	62	44
19	37.sff	М	33	24	61	11	19	73
20	12.sff	М	15	39	63	26	35	58
21	03.sff	Т	35	68	16	41	56	24
22	11.sff	Т	28	74	19	40	53	15
23	26.sff	Т-БД	15	93	8	18	95	10
24	28.sff	Т	61	37	12	55	48	18
25	41.sff	Т	35	67	10	31	57	15
26	43.sff	Т	42	69	9	29	58	18
27	44.sff	Т	45	58	14	37	64	21
28	48.sff	Т	20	55	19	29	55	20
29	52.sff	Т	23	72	16	25	68	17
30	58.sff	Т	34	63	15	38	71	19
31	59.sff	Т	59	44	18	68	40	12
32	64.sff	Т	15	59	20	25	60	18

Продовження 2 Таблиці 2.5. Результати класифікації образів підводних об'єктів

№ п/п	Ім'я файла	Експертна оцінка	Ймовірність збігу з образами БД, %					
			по коду строки			по коду стовпця		
			С	Т	М	С	Т	М
33	66.sff	Т	28	60	22	32	58	10
34	67.sff	Т	39	57	17	22	61	12
35	75.sff	Т	25	74	11	40	55	16
36	78.sff	Т	29	76	12	36	59	21
37	80.sff	Т	34	62	13	31	65	19
38	94.sff	Т	18	69	15	43	73	15
39	97.sff	Т	65	38	15	48	55	18
40	10.sff	С	72	36	19	80	24	9
41	21.sff	С	75	32	16	81	20	14
42	22.sff	С	63	49	13	57	44	8
43	32.sff	С	58	40	26	69	36	10
44	50.sff	С-БД	97	14	7	94	18	11
45	05.sff	С	54	60	19	49	68	15
46	09.sff	С	63	27	13	71	34	18
47	81.sff	С	65	28	15	76	38	17
48	89.sff	С	58	35	18	68	29	15
49	20.sff	С	72	39	23	44	40	57
50	24.sff	С	79	30	21	64	35	22
51	29.sff	С	75	29	9	73	33	15
52	39.sff	С	68	34	11	71	27	17
53	40.sff	С	55	40	29	59	43	12
54	95.sff	С	43	58	19	64	39	10
55	51.sff	С	69	38	21	79	33	7
56	57.sff	С	64	35	16	65	37	15

Продовження 3 Таблиці 2.5. Результати класифікації образів підводних об'єктів

№ п/п	Ім'я файла	Експертна оцінка	Ймовірність збігу з образами БД, %					
			по коду строки			по коду стовпця		
			С	Т	М	С	Т	М
57	73.sff	С	59	32	10	72	35	18
58	90.sff	С	53	41	13	42	63	21

Примітка: М - «Мина», Т - «Торпеда», С - «Судно», М-БД - «Мина» з БД, Т-БД - «Торпеда» з БД, З-БД - «Судно» з БД. Жирним шрифтом виділені найбільші значення ймовірності збігу з образами БД, помилки виділені сірим.

В результаті аналізу даних таблиці 3 встановлено наступне:

- 1) ймовірність розпізнавання невідомих підводних об'єктів в середньому склала більше 80% при БД в три образи;
- 2) ймовірність розпізнавання об'єктів у БД склала більше 85%;
- 3) помилка при розпізнаванні об'єкта типу «Мина» склала 20%, помилка при розпізнаванні об'єкта типу «Торпеда» - 16%, помилка при розпізнаванні об'єкта типу «Судно» - 21%.

Порівнюючи описаний метод розпізнавання образів підводних об'єктів (ймовірність розпізнавання більше 80%) з відомим в літературі методом (ймовірність розпізнавання не більше 70%), заснованим на аналізі ехосигналів, отриманих від підводних об'єктів, за найбільшим значенням відгуку узгодженого фільтра, який реалізує взаємно кореляційну обробку сигналу і зондуючого імпульсу, слід зазначити перевага першого.

2.2.3. Підсумок за ідентифікаційним методом розпізнавання образів малорозмірних підводних об'єктів.

Головними достоїнствами запропонованого ідентифікаційного методу є висока чутливість, відсутність помітних крайових ефектів і, отже,

можливість виявляти об'єкти при малих ЗСШ. У ході дослідження підтвердилася можливість застосування ідентифікаційного методу для розпізнавання образів підводних об'єктів, шляхом аналізу відбитого луна-сигналу.

Розрахункова ймовірність виявлення об'єкта на тлі перешкод за допомогою описаного ідентифікаційного методу склала понад 0,97. Ймовірність розпізнавання і класифікації підводних об'єктів визначається повнотою заздалегідь створеної бази образів відомих об'єктів, сформованої відповідно до оцінок сертифікованих експертів. При базі даних в три образи ймовірність розпізнавання склала більше 80%.

Практична значимість отриманих результатів полягає в можливості реалізації для ГБЗ принципово нового типу програмного забезпечення, яке дозволить підвищити їх інтелектуалізацію в області розпізнавання і класифікації підводних об'єктів.

Подання та опис об'єктів або областей, виділених при сегментації зображення, складають, як почати роботу більшості автоматизованих процесів, в яких беруть участь цифрові зображення. Такі описи, наприклад, є вхідною інформацією для методів розпізнавання об'єктів, що обговорюються в наступному розділі. Як показує викладений вище набір методів опису, вибір того чи іншого підходу визначається розглянутої завданням. Мета полягає в тому, щоб вибрати дескриптори, здатні «схопити» істотні відмінності між об'єктами або класами об'єктів, і які в той же час були б максимально незалежними від змін таких факторів, як розташування об'єкта, його розмір і орієнтація.

З АНАЛІЗ ТА КЛАСИФІКАЦІЯ АЛГОРИТМІВ СЕГМЕНТАЦІЇ ЗОБРАЖЕНЬ

Розглянемо класифікації алгоритмів сегментації зображень і які вимоги формуються до нових алгоритмів. Для розробки нових алгоритмів пропонується триетапна структурна схема, як версії реалізації якої обговорюється алгоритм квазіоптимальний сегментації. Вводиться ознака, що відображає особливості групи ієрархічних алгоритмів, до якої належить запропонований алгоритм. Введення нової ознаки дозволяє уточнити схему класифікації і вимоги до створюваних алгоритмів.

Класифікації алгоритмів сегментації зображень (АСЗ) присвячений ряд досліджень [1-13]. Безліч схем класифікації АСЗ умовно розділяється на наступні категорії:

- змішані класифікації;
- спеціалізовані класифікації;
- класифікації за єдиною ознакою;
- узагальнені класифікації.

Особливий інтерес представляють схеми узагальненої класифікації, оскільки вони, по-перше, дозволяють однозначно класифікувати існуючі АСЗ, по-друге, передбачають появу нових АСЗ, які виявляються в схемі по "білих плям", і, по-третє, формують вимоги до нових АСЗ.

В якості нового алгоритму, вимоги до якого викладені в роботі [10], у цій статті розглядається алгоритм ієрархічної сегментації зображень, на виході якого формується безліч розбиття вихідного зображення на кластери пікселів з числом від 1 до N , де N - число пікселів в зображенні. Кожне наступне розбиття виходить з попереднього послідовним поділом кластера (сегмента) або об'єднанням пари пікселів, сегментів або кластерів. Значення пікселів всередині кожного кластера (сегмента) усереднені за кольором або за яскравістю в разі зображень з відтінків сірого. Тому кожне розбиття зображення кусково-постійно.

Генерація оптимальних наближень є обчислювально складним завданням. Обчислювальну складність можна знизити за рахунок побудови ієрархічної структури даних і застосування ряду допоміжних алгоритмів поліпшення якості. Оскільки ієрархічні розбиття зазвичай істотно відрізняються від оптимальних, застосовується процедура їх оптимізації. На виході АСЗ формуються Квазіоптимальний наближення, тобто розбиття, близькі до оптимальних.

Розглянемо такі теми:

- розвиток схеми узагальнення класифікації;
- визначення нової класифікаційної ознаки на основі властивості, характерного для груп ітераційних алгоритмів і алгоритмів з ієрархічною структурою даних;
- уточнення схеми класифікації алгоритмів сегментації [10] за рахунок введення додаткового класифікаційної ознаки;
- демонстрації переваг алгоритму квазіоптимальної сегментації.

3.1. Узагальнена класифікація алгоритмів сегментації зображень

Ідея об'єднання кількох класифікацій з єдиними ознаками в одну схему висунута в роботі [2]. На рис. 3.1 представлена схема узагальненої на основі ознак першого рівня класифікації АСЗ.

Також відзначемо доцільність декомпозиції запропонованої схеми класифікації. Її подальша деталізація виконана в роботі [10], де представлений також огляд схем класифікацій АСЗ. Класифікаційні ознаки другого рівня - тип зображення і наявність критерію - показані на рис. 3.2.

Дворівнева узагальнена схема класифікації АСЗ, відповідно до результатів огляду, наведена на рис. 3.3, де фоном виділені сектори, для яких існує хоча б один алгоритм сегментації; для білих секторів ("білі плями") не знайдено жодного з розглянутих алгоритмів (рис. 3.3.а).

		Способ обработки	
		Выделение границ (разрывность)	Выделение областей (сходство)
Стратегия исполнения	Параллельные вычисления	G1: алгоритмы параллельной обработки на основе выделения границ	G3: алгоритмы параллельной обработки на основе выделения областей
	Последовательные вычисления	G2: алгоритмы последовательной обработки на основе выделения границ	G4: алгоритмы последовательной обработки на основе выделения областей

Рисунок 3.1 – Схема узагальнення на основі ознак першого рівня

		Наличие критерия качества	
		С установленным критерием качества	Без установленного критерия качества
Тип изоб- раже- ния	Цветные изображения	W1: алгоритмы сегментации цветных изображений с использованием критерия качества	W3: алгоритмы сегментации цветных изображений без использования критерия качества критерия
	Полутоновые изображения	W2: алгоритмы сегментации полутоновых изображений с использованием критерия качества	W4: алгоритмы сегментации полутоновых изображений без использования критерия качества

Рисунок 3.2 – Класифікаційні ознаки другого рівня - тип зображення і наявність критерію

Відповідно до цієї схеми (рис. 3.3.б) актуальна розробка алгоритмів:

- виділення меж з використанням критерію якості (наприклад, мінімум сумарної квадратичної помилки) для будь-якого типу зображень і будь-якої стратегії обчислень (сектори G1_W1, G1_W2, G2_W1, G2_W2);
- виділення областей напівтонових зображень з використанням критерію якості і послідовним виконанням операцій (сектор G4_W2).

Можна зазначити обґрунтованість значущості алгоритму квазіоптимальний сегментації, що відповідає класифікаційними ознаками сектора G4_W2 (якому відповідає "біла пляма", рис. 3.3.а). Оскільки напівтонові зображення обробляються як окремий випадок кольорових, то розглянутий АСЗ відповідає також і класифікаційними ознаками G4_W1.

		Выделение границ		Выделение областей	
		С установленным критерием качества	Без установленного критерия качества	С установленным критерием качества	Без установленного критерия качества
Параллельное вычисление	Цветное изображение	W1	W3	W1	W3
	Полутонное изображение	W2	W4	W2	W4
Последовательное вычисление	Цветное изображение	W1	W3	W1	W3
	Полутонное изображение	W2	W4	W2	W4

а)

		Выделение границ		Выделение областей	
		С установленным критерием качества	Без установленного критерия качества	С установленным критерием качества	Без установленного критерия качества
Параллельное вычисление	Цветное изображение	W1	W3	W1	W3
	Полутонное изображение	W2	W4	W2	W4
Последовательное вычисление	Цветное изображение	W1	W3	W1	W3
	Полутонное изображение	W2	W4	W2	W4

б)

Рисунок 3.3 – Класифікаційні ознаками сектора

3.2 Алгоритм квазіоптимальний сегментації

Розглянемо алгоритм сегментації цифрових зображень, структурна схема якого представлена на рис. 3.4. Алгоритм сегментації генерує ієрархічну послідовність розбиття вихідного зображення на кластери від 1 до N. Кожне розбиття являє собою сукупність кусочно-постійних областей. Приклад вихідний послідовності розбиттів показаний на рис. 3.5.

В основі розглянутого в даній роботі алгоритму квазіоптимальний сегментації лежить метод Уорда. Метод полягає в ітеративном поділі пікселів на кластери або, навпаки, їх злиття в один кластер. Пікселі, розділені

на безлічі, усереднені за значенням яскравості всередині кожної множини. Класичний метод Уорда має ряд істотних недоліків:

- обчислювальна складність методу квадратично зростає з лінійним зростанням числа елементів (пікселів зображення), що перешкоджає безпосередньому застосуванню методу при цифровій обробці зображень;
- метод нестійкий щодо вибору варіанту в разі рівнозначних рішень; на перших кроках укрупнення груп елементів (пікселів, сегментів, кластерів) зустрічаються пари, операції над якими відповідають однаковим значенням цільової показовою функції.

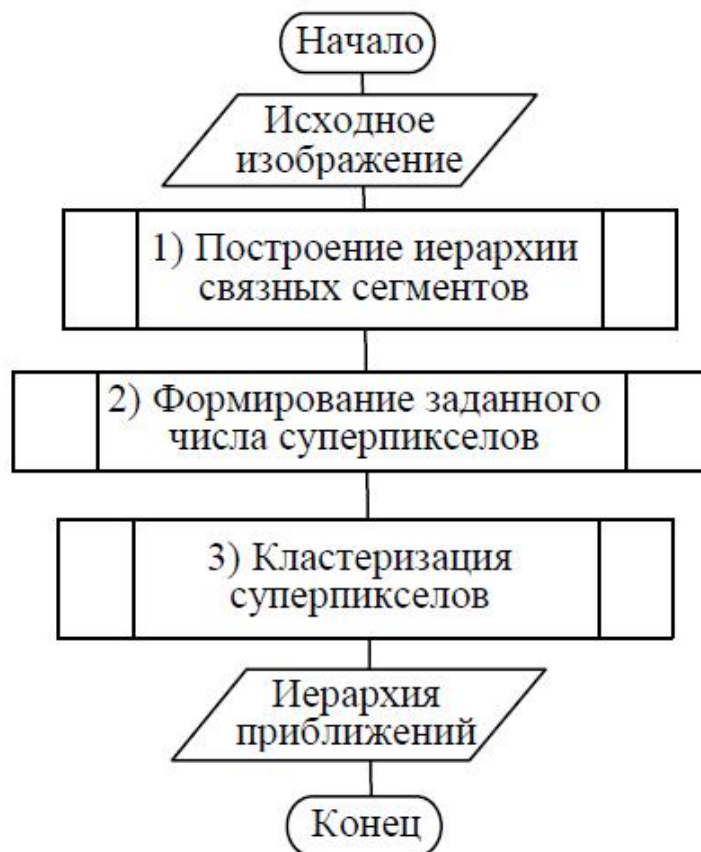


Рисунок 3.4 – Структурна схема алгоритму сегментації цифрових зображень

Розглянутий алгоритм квазіоптимальний сегментації забезпечує зниження обчислювальної складності за рахунок поділу процесу побудови ієрархічної послідовності розбиттів на три етапи (рис. 3.4).

1-й етап - швидке побудова ієрархічної послідовності грубого розбиття. Швидка побудова здійснюється за рахунок зменшення числа пар пікселів (сегментів, кластерів) на кожній ітерації. Це призводить до неминучого зниження якості. Редукція числа розглянутих пар до тільки суміжних зводить метод Уорда до моделі Мамфорда - Шаха. На цьому етапі будується ієрархія зв'язкових сегментів у спрощеній моделі Мамфорда - Шаха (без урахування довжини кордону сегмента). Для побудови ієрархії застосовується операція Merge злиття двох кластерів.

2-й етап - проміжне поліпшення якості розбиття при фіксованому числі кластерів пікселів (квітів). Для цього виконуються наступні операції:

- Divide - поділ кластера надвоє (до розбиття на попередньому кроці);
- Split - виділення частини пікселів в окремий кластер (частина пікселів утворює окремий самостійний кластер);
- Correct - рекласифікація частини пікселів кластера (частина пікселів переноситься в інший існуючий кластер).

На цьому етапі реалізується спеціально розроблений SI-метод (Segmentation Improvement) поліпшення якості сегментації, який описаний в роботі [12]. SI-метод полягає в циклічному поділі надвоє сегмента в одному місці зображення і злиття пари суміжних сегментів в іншому місці за критерієм зменшення значення функціоналу якості для даного розбиття. Результат виконання другого етапу - формування суперпікселів (укрупнених сегментів).

3-й етап - "добудова" ієрархічної послідовності розбиттів класичним методом Уорда. На вході трьохетапного алгоритму квазіоптимальний сегментації, представленого на рис. 3.5, вихідне зображення півтонове або кольорове. У програмі передбачений параметр - число суперпікселів, який задає "точність" обчислень і управляє процесом сегментації зображень. Цей параметр дозволяє варіювати розбиття між чистою сегментацією - грубою, але швидкою за версією моделі Мамфорда - Шаха, і точною, але повільною кластеризацією пікселів по методу Уорда.

На виході алгоритм генерує безліч розбиття (рис. 3.5). Кожне розбиття з послідовності характеризується значенням функціоналу якості (сумарною квадратичною помилки E). Тестове зображення, результат обробки якого представлений на рис. 3.5. На його прикладі в роботах виконано зіставлення двох методів автоматичної сегментації зображень: розглянутого в цій статті алгоритму квазіоптимальної сегментації і методу сегментації за кольором і текстурою. На рис. 3.5 представлені результати розбиття зображень - перші шість наближень і 20000-е з можливих 273056. Над кожною "картинкою" розбиття зазначено число кластерів (квітів), а під "картинками" - відповідне значення сумарної квадратичної помилки.

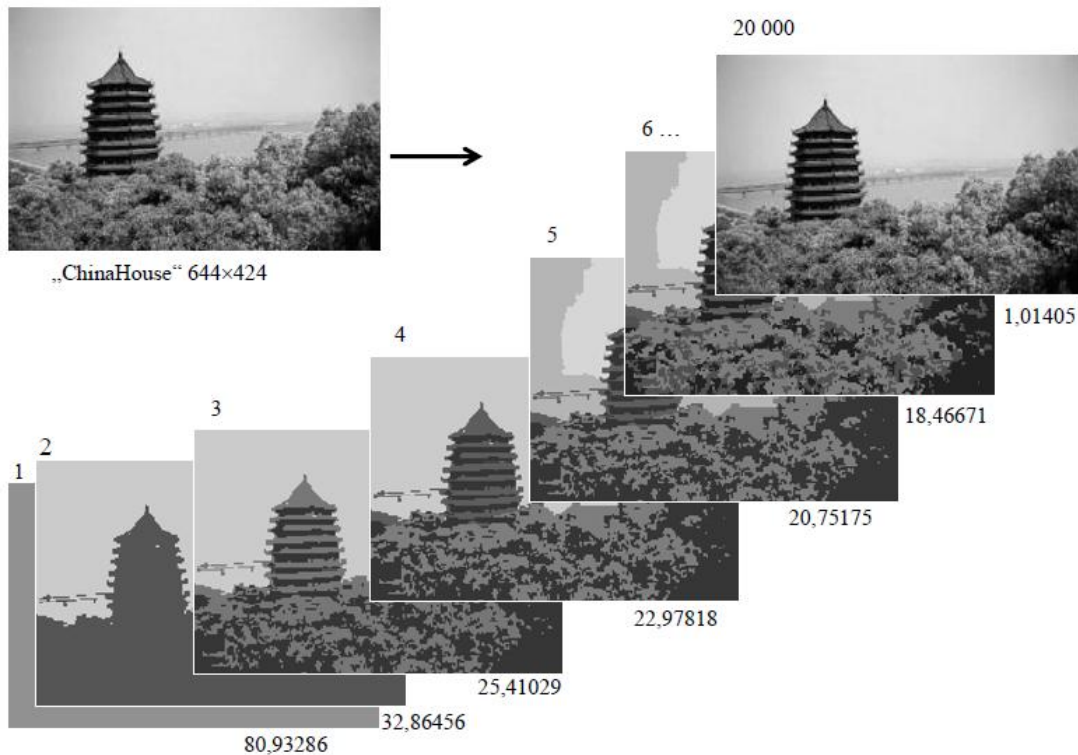


Рисунок 3.5 – На вході трьохетапного алгоритму квазіоптимальний сегментації

Алгоритм квазіоптимальної сегментації призначений для виділення областей напівтонових і кольорових зображень і використовує критерій

якості при послідовному формуванні ієрархії наближень. В узагальненій схемі класифікації АСЗ, запропонованої в роботі [10], що розглядається алгоритм відноситься до секторів G4_W1 і G4_W2. Таким чином, алгоритм забезпечує "закриття" "білої плями".

3.3 Класифікації за п'ятьма ознаками

Схема класифікації не враховує істотний класифікаційний ознака групи до якої належить розглянутий алгоритм квазіоптимальний сегментації, що генерує безліч розбиття аналогічно методу пірамідальної сегментації. У зв'язку з цим набір чотирьох ознак доцільно розширити до п'яти, ввівши ознака числа розбиття на виході АСЗ: єдине розбиття / безліч розбиття. Ознака відображає як вихідну характеристику АСЗ, так і його внутрішню особливість, вказуючи на фіксований або змінюваний число розглядає ваємих однорідних множин (кластерів, сегментів, табл. 3.1).

Таблиця 3.1 Значення класифікаційних ознак

Класифікаційний ознака	Значення ознаки
Тип операції	Виділення областей зображення на основі властивостей подібності
Стратегія обчислень	Послідовне обчислення наближень
Наявність критерію	Функціонал якості - сумарна квадратична помилка; критерій якості - мінімум збільшення помилки
Тип зображення	Кольорове і півтонове
Число розбиття вихідного зображення на виході алгоритму	Безліч розбиття зображення за кількістю розглянутих кластерів, число кластерів змінюється від 1 до N

У табл. 3.1 наведені значення класифікаційних ознак, що характеризують розглянутий алгоритм ієрархічний сегментації, на виході якого формується безліч квазіоптимальних розбиття вихідного зображення на кластери від 1 до N (рис. 3.6).

		Выделение границ		Выделение областей	
		С установленным критерием качества	Без установленного критерия качества	С установленным критерием качества	Без установленного критерия качества
Параллельное вычисление	Цветное изображение	1 M	1 W1 W3 M	1 M	1 W1 W3 M
	Полутонное изображение	1 M	1 W2 W4 M	1 M	1 W2 W4 M
Последовательное вычисление	Цветное изображение	1 M	1 W1 W3 M	1 M	1 W1 W3 M
	Полутонное изображение	1 M	1 W2 W4 M	1 M	1 W2 W4 M

а)

		Выделение границ		Выделение областей	
		С установленным критерием качества	Без установленного критерия качества	С установленным критерием качества	Без установленного критерия качества
Параллельное вычисление	Цветное изображение	1 M	1 W1 W3 M	1 M	1 W1 W3 M
	Полутонное изображение	1 M	1 W2 W4 M	1 M	1 W2 W4 M
Последовательное вычисление	Цветное изображение	1 M	1 W1 W3 M	1 M	1 W1 W3 M
	Полутонное изображение	1 M	1 W2 W4 M	1 M	1 W2 W4 M

б)

Рисунок 3.6 – Моделі: а) схема узагальненої класифікації за п'ятьма ознаками, б) світло-сірим тоном відзначено місце в схемі розглянутого алгоритму квазіоптимальної сегментації (G4_W1_M, G4_W2_M); на схемі М - множина розбиття, що генерується на виході АСЗ, "1" - єдине розбиття

Алгоритми, що генерують безліч розбиття, мають перевагу щодо алгоритмів, що генерують єдине розбиття. З безлічі розбиття вибирається одне найбільш підходяще. Вибір здійснює або користувач - "фахівець-

предметник", або автоматизована система обробки зображень. Вибір проводиться на етапі постобробки, наступного за етапом сегментації зображення. Відповідне розбиття вибирається в умовах конкретного завдання, при цьому враховуються завжди апріорна інформація про тематику сцени, особливості зйомки, об'єкти інтересу, їх кількість, особливості і характеристики.

3.4 Аналіз класифікаційних ознак

У зведеній рис. 3.7 наведені класифікаційні ознаки і їх короткі характеристики. Таблицю утворюють чотири графи: "Автор", "Класифікаційна ознака", "Значення ознаки" і "Примітка". Тут варто пояснити, що в графі "Значення ознаки" наведені обидва значення відповідного класифікаційної ознаки, однак ознака приймає тільки одне значення; в графі "Примітка" наводяться особливості реалізації алгоритму при обраному значенні ознаки.

Два класифікаційних ознаки формують рівень. Так, в роботі [11] вперше сформована однорівнева схема класифікації АСЗ з перших двох класифікаційних ознак, раніше запропонованих в [2]. В роботі [10] запропоновано третій і четвертий класифікаційні ознаки, з яких сформований другий рівень. Ознаки обох рівнів об'єднані в одну дворівневу схему класифікації АСЗ (див. Рис. 3.3). Введений в даній роботі ще один класифікаційний ознака дозволяє після декомпозиції дворівневої схеми класифікації АСЗ виявити раніше приховані "білі плями". Новий класифікаційний ознака об'єднує алгоритми, що генерують безліч розбиття на виході, в окрему групу.

Для подальшого дослідження представляється перспективним:

- виключити класифікаційний ознака "тип зображення"; вважати зображення з відтінків сірого окремими випадками кольорових;
- ввести шосту ознаку - "наявність ієрархічної структури даних".

Автор	Класифікаційний признак	Значение признака	Примечание
Гонзалес — Вудс [7]	1. Способ обработки	Нахождение области	Выделение объекта по свойствам схожести
		Нахождение границ	Выделение объекта по свойствам различия
Розенфельд [8]	2. Стратегия исполнения	Последовательные вычисления	Последовательное исполнение вычислительных операций
		Параллельные вычисления	Параллельное исполнение вычислительных операций
Поршнеv [10]	3. Тип изображения	Полутоновое изображение	Алгоритмы сегментации изображений с одной яркостной составляющей
		Цветное изображение	Алгоритмы сегментации, учитывающие все три цветовые составляющие изображения
	4. Наличие критерия качества	С критерием качества	Алгоритмы сегментации с использованием установленного критерия качества
		Без критерия качества	Алгоритмы сегментации без использования установленного критерия качества
Ханыков (настоящая статья)	5. Число разбиений исходного изображения на выходе алгоритма	Единственное разбиение	Алгоритм генерирует на выходе одно отсегментированное на однородные области разбиение исходного изображения. В ходе вычислительного процесса число однородных по некой характеристике множеств (сегментов, кластеров) фиксировано
		Множество разбиений	Алгоритм генерирует на выходе множество отсегментированных на однородные области разбиений исходного изображения. В ходе вычислительного процесса число однородных по некой характеристике множеств (сегментов, кластеров) варьируется в некотором диапазоне

Рисунок 3.7 – Схема класифікації ознак АСЗ

Ієрархічна структура даних дозволяє повернутися до попереднього стану (розбиття) або знайти нове розбиття, краще за значенням критерію якості. Ознака "наявність ієрархічної структури даних" взаємопов'язаний з п'ятою ознакою - "число розбиття на виході АСЗ". Наявність вбудованої ієрархічної структури даних однозначно тягне генерацію множини розбиття на виході АСЗ.

Видається менш перспективним пошук і розробка алгоритмів виділення кордонів з використанням критерію якості (сектори G1_W1, ..., G2_W2 на рис. 3.6). Це завдання успішно вирішене за допомогою послідовності алгоритмів виділення областей з використанням встановленого критерію якості і алгоритмів виділення кордонів без використання критерію

якості (наприклад, фільтр Лапласа, фільтр Робертса та ін.. Перші формують гомогенні області, на кордонах яких утворюються різкі перепади за кольором (або відтінкам сірого), другі алгоритми виконують фільтрацію, за результатами якої виділяється межа об'єкта.

Керуючись схемою, наведеною на рис. 3.6. в подальшому передбачається розвивати алгоритми виділення областей з використанням встановленого критерію якості та ієрархічною структурою даних.

ВИСНОВКИ

Наведений огляд методів пошуку малорозмірних об'єктів на зображенні дозволяє зробити наступні висновки. Існуючі в даний час алгоритми сегментації малорозмірних об'єктів адаптовані до роботи із спеціалізаціями де потрібні жорстко спрямованої дії, тобто призначені для ефективного вирішення завдання в конкретній задачі із рядом обмежувальних умов.

У даній роботі проаналізовано та класифіковано АСЗ, що передбачає напрямок перспективних досліджень - розробку нових алгоритмів комп'ютерного зору.

Запропоновано структурну схему алгоритмів, як приклад реалізації якої представлений алгоритм квазіоптимальної сегментації, що генерує безліч розбиття вихідного зображення по числу розглянутих кластерів.

Введено додаткову класифікаційну ознаку (число розбиття на виході АСЗ) і виконана декомпозиція узагальненої схеми класифікації АСЗ до п'яти ознак.

Представляється перспективним виняток ознаки "тип зображення" і введення нової ознаки "наявність ієрархічної структури даних" з подальшою ревізією всіх АСЗ.

Надалі передбачається розробка алгоритмів виділення областей з використанням встановленого критерію якості та ієрархічною структурою даних, що забезпечують швидкі операції над множинами пікселів.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Андреев, В. В. Как организовать делопроизводство на предприятии [Текст] / В. В. Андреев. – М. : ИНФРА-М, 1997. – 94 с.
2. Gonzalez R., Woods R. Digital Image Processing. NJ: Prentice Hall, 2002
3. Фісенко В. Т., Вілес Л. Д., Можейко В. І., Фісенко Т. Ю. Виявлення малорозмірних об'єктів на підставі багатоцільового супроводу по послідовності відеокадрів // Зб. тр. міжнар. конф. «Прикладна оптика-2012». СПб, 2012. Т. 1. С. 118-122.
4. Фісенко В. Т., Вілес Л. Д., Можейко В. І., Фісенко Т. Ю. Про оцінку координат міжкадрових зсувів в відеосистемі кругового огляду. // Зб. доп. XV міжнар. конф. по м'яких обчислень та вимірювань. СПб, 2012. Т. 1. С. 149-152.
5. Фісенко В. Т., Можейко В. І., Фісенко Т. Ю. Автоматичний супровід об'єктів в комп'ютерних системах обробки зображень // Оптичний журнал. 2007. Т. 74, № 11. С. 39-46.
6. Hyeonwoo Cho, Jeonghwe Gu, Son-Cheol Yu. Robust Sonar-Based Underwater Object Recognition Against Angle of-View Variation // Sensors Journal IEEE. 2016. Vol. 16, no. 4. P. 1013-1025.
7. Vandrish P., Vardy A., Walker D., Dobre O. A. Side-scan sonar image registration for AUV navigation // Underwater Technology (UT) // IEEE Symposium on and 2011 Workshop on Scientific Use of Submarine Cables and Related Technologies (SSC). 2011. P. 1-7.
8. Bianco G., Gallo A., Bruno F., Muzzupappa M. A comparison between active and passive techniques for underwater 3D applications // International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. ISPRS Archives. 2011. Vol. 38 (5W16). P. 357-363.
9. Azimi-Sadjadi M.R., Yao D., Huang Q., Dobeck G. J. Underwater target

classification using wavelet packets and neural networks // IEEE Transactions on Neural Networks. 2000. Vol. 11, no. 3. P. 784-794.

10. Поршневу С. В., Левашкіна А. О. Універсальна класифікація алгоритмів сегментації зображень // Журн. наукових публікацій аспірантів і докторантів. 2008. № 3 (31). С. 163-172.

11. Semantic-Based Visual Information Retrieval / Ed. Y. J. Zhang. IGI Global, 2006.

12. Харін М. В., Хаников І. Г. Комбінований метод поліпшення сегментації зображення // Тр. Бурятського держ. ун-ту. 2015. С. 118-124.