

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

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

Кафедра електронних обчислювальних машин
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА
Пояснювальна записка

Рівень вищої освіти другий (магістерський)

Методи класифікації та кластеризації даних
з використанням машинного навчання

(тема)

Виконав:

студент II курсу, групи СПМ-21-2
Щепка О.О.
(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. Міхаль О.П.
(посада, прізвище, ініціали)

Допускається до захисту

Зав. кафедри ЕОМ

(підпис)

Коваленко А.А.

(прізвище, ініціали)

2023 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 КВАЛІФІКАЦІЙНУ РОБОТУ

студенту Щепці Олексію Олександровичу
(прізвище, ім'я, по батькові)

1. Тема роботи Методи класифікації та кластерізації даних з використанням машинного навчання

затверджена наказом по університету від “ 03 ” квітня 2023 р. № 318 Ст

2. Термін подання студентом роботи до екзаменаційної комісії 17 травня 2023 р.

3. Вхідні дані до роботи _____

машинне навчання

штучні нейронні мережі

класифікація

кластерізація

4. Перелік питань, що потрібно опрацювати у роботі _____

Аналіз предметної області

Методи машинного навчання для сегментації та локалізації зображень

Реалізація та тестування моделі

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій (слайдів) 13 слайдів

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	Отримання завдання. Аналіз предметної області	03.04.2023–10.04.2023	
2	Аналіз існуючих моделей та методів	11.04.2023–26.04.2023	
3	Розробка моделі	27.04.2023–29.04.2023	
4	Аналіз штучних нейронних мереж	30.04.2023–02.05.2023	
5	Отримання та аналіз результатів	03.05.2023–06.05.2023	
6	Оформлення пояснювальної записки	07.05.2023–13.05.2023	

Дата видачі завдання 03 квітня 2023 р.

Студент _____
(підпис)

Керівник роботи _____
(підпис)

проф. Міхаль О.П.
(посада, прізвище, ініціали)

РЕФЕРАТ

Пояснювальна записка кваліфікаційної роботи: 74 с., 20 рис., 1 дод., 12 джерел.

МАШИННЕ НАВЧАННЯ, КЛАСИФІКАЦІЯ, КЛАСТЕРИЗАЦІЯ,
ШТУЧНА НЕЙРОННА МЕРЕЖА, ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ.

Метою кваліфікаційної роботи є аналіз існуючих методів класифікації та кластеризації даних з використанням машинного навчання.

У ході виконання кваліфікаційної роботи Проведено аналіз існуючих методів класифікації та кластеризації даних з використанням машинного навчання. аналіз існуючих методів класифікації та кластеризації даних. Проведено аналіз існуючих штучних нейронних мереж.

Відповідно завданню проведено аналіз сучасного стану питання, досліджені існуючі моделі і підходи до сегментації і локалізації об'єкту на цифровому зображенні з використанням методів машинного навчання. Побудовано відповідну загальну модель локалізації, з використанням якої проведений експеримент.

ABSTRACT

Master's thesis: 74 pages, 20 figures, 1 appendices, 12 sources.

MACHINE LEARNING, CLASSIFICATION, CLUSTERIZATION,
ARTIFICIAL NEURAL NETWORK, SOFTWARE.

The major goal of this thesis is to analyze the existing methods of data classification and clustering using machine learning.

In order to analysis of existing methods of data classification and clustering using machine learning was carried out. analysis of existing data classification and clustering methods. An analysis of existing artificial neural networks was carried out. In accordance with the task, an analysis of the current state of the issue was carried out, existing models and approaches to object segmentation and localization on a digital image were investigated using machine learning methods. An appropriate general localization model was constructed, using which the experiment was conducted.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ	7
ВСТУП	8
1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ	10
1.1 Завдання і проблеми	10
1.2 Сегментація зображень та методи машинного навчання	18
1.3 Методи сегментації границь зображень	24
1.4 Методи нарощування областей	32
1.5 Текстура як цінна ознака якості зображення	34
1.6 Сегментація зображень на основі класичних методів.....	35
1.7 Сегментація зображень на основі кластерного аналізу даних та штучних нейронних мереж	38
1.8 Сучасний стан і перспективи розвитку методів сегментації.....	42
1.9 Локалізація зображень.....	42
2 МЕТОДИ МАШИННОГО НАВЧАННЯ ДЛЯ СЕГМЕНТАЦІЇ ТА ЛОКАЛІЗАЦІЇ ЗОБРАЖЕНЬ.....	43
2.1 Класична модель сегментації.....	43
2.2 Виділення границь за допомогою операторів.....	47
2.3 Розробка алгоритму сегментації зображення.....	52
2.4 Локалізація з використанням методу головних компонент	55
3 РЕАЛІЗАЦІЯ ТА ТЕСТУВАННЯ МОДЕЛІ	61
3.1 Загальна модель сегментації і локалізації об'єктів	61
ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ	65
ДОДАТОК А Графічний матеріал кваліфікаційної роботи.....	67

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ
І ТЕРМІНІВ

КК – карти Кохонена

МН – машинне навчання

ПЗ – програмне забезпечення

СТЗ – система технічного зору

ШНМ – штучна нейронна мережа

ВСТУП

Останнім часом з'явилась необхідність розпізнавання будь-яких об'єктів на зображеннях, зокрема ідентифікації людей. Для розпізнавання певних груп осіб та надання їм доступу, на підприємствах встановлюють електронні турнікети з ідентифікацією відбитків пальців або зі зчитувачами карток, але на сьогоднішній день, у зв'язку з бурхливим розвитком біометричних технологій, компанії все частіше переходять на інші методи розпізнавання, більш точні та зручні. Наприклад, розпізнавання осіб у відеопотоці в режимі реального часу. Поряд з розвитком методів розпізнавання та стрімкого зростання застосування біометричних систем, методи ідентифікації по обличчю виявляють свої недоліки та стають уразливими. Але, наприклад, в епідеміологічних ситуаціях, таких як пандемія коронавірусу COVID-19, захисні медичні маски стали необхідністю для людей, що у громадських місцях.

Маски є захисним засобом, який використовували люди у повсякденному житті для захисту від коронавірусу. Така ситуація робить звичайну технологію контролю доступу неефективною у багатьох випадках, а саме таких як контроль доступу в приміщеннях, контроль доступу з використанням ідентифікації по особі, підрахунок відвідуваності, перевірка безпеки на вокзалах та ін. У зв'язку з великим попитом до технологій розпізнавання та ідентифікації, особливо в сьогоднішні дні, виникла велика необхідність поліпшення методів розпізнавання для існуючих способів технології розпізнавання осіб. Багато сучасних розробників використовують методи машинного навчання для розв'язань завдань класифікації та кластеризації даних, які задані неявно.

Машинне навчання – це набір алгоритмів та інструментів, які допомагають машинам розуміти закономірності в даних та використовувати цю базову структуру для виконання дії та прийняття рішення про поставленої

задачі. Є багато способів, якими машини прагнуть зрозуміти закономірності, що у основі даних.

Метою кваліфікаційної роботи є аналіз існуючих методів класифікації та кластеризації даних з використанням машинного навчання.

В ході виконання кваліфікаційної роботи повинні бути виконані наступні завдання:

- аналіз існуючих методів класифікації та кластеризації даних;
- аналіз існуючих методів машинного навчання;
- аналіз алгоритмів та моделей локалізації та сегментації зображень;
- розробка моделі локалізації об'єктів;
- вибір типу штучної нейронної мережі;
- реалізація та тестування моделі.

1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ

1.1 Завдання і проблеми

Потужний розвиток засобів комп'ютерної техніки послужив поштовхом для напрямку, що пов'язаний з розробкою теоретичних основ і практичною реалізацією систем, призначених для розпізнавання та обробки зображень. Принциповими питаннями в теорії обробки зображень є питання: формування, введення, подання в комп'ютері й візуального відтворення зображення. Засоби обробки зображень в даний час одержують усе більше поширення, і важко назвати таку галузь науки чи сферу виробничої діяльності, де вони не використовуються чи не будуть використовуватися в найближчі роки.

Складовою частиною систем розпізнавання та промислових роботів є обробка зображень. Обробка зображень з метою їх розпізнавання є однією з центральних і практично важливих завдань під час створення систем штучного інтелекту. Проблема створення систем штучного інтелекту носить явно виражений комплексний ієрархічний характер і включає ряд основних етапів: сприйняття поля зору, сегментація зображення, нормалізація виділених об'єктів, розпізнавання. Такий важливий обов'язковий етап як розуміння (інтерпретація) зображень включається частково в етап сегментації і остаточно вирішується на етапі розпізнавання.

Сегментацією зображення називається розбиття зображення на несхожі за деякою ознакою області. Передбачається, що області відповідають реальним об'єктам або їх частинам, а межі областей відповідають кордонам об'єктів. Сегментація відіграє важливу роль в задачах обробки зображень та комп'ютерного зору. Сегментація необхідна при розпізнаванні образів та аналізі сцен, оскільки контури, як правило, є найбільш інформативними й ненадлишковими ознаками зображення, що обробляється. Для проведення

сегментації зображень найбільш широке застосування одержали методи, засновані на різного роду статистичних і імовірнісних моделях, методах кластерного аналізу, порогових та градієнтних методах, теорії графів.

У комп'ютерному зорі сегментація - це процес поділу цифрового зображення на кілька сегментів (множина пікселів, також званих суперпікселями). Мета сегментації полягає у спрощенні та зміні подання зображення, щоб його було простіше і легше аналізувати. Результатом сегментації зображення є множина сегментів, які разом покривають все зображення, або множина контурів, виділених із зображення. Всі пікселі в сегменті схожі за деякою характеристикою або за обчисленою властивістю, наприклад за кольором, яскравості або текстурі. Сусідні сегменти значно відрізняються по цій характеристиці.

Сегментація зазвичай розуміється як процес пошуку однорідних областей на зображенні. Цей етап дуже важкий і в загальному вигляді не алгоритмізований до кінця для довільних зображень. Зауважимо, що один з найефективніших методів нарощування областей передбачає вибір стартових точок або за допомогою оператора (алгоритм центроїдного зв'язування), або автоматично. Ефективним тут є метод вододілів з подальшою угрупованням навколо них областей по зв'язності.

Всі методи дуже прийнятні з точки зору обчислювальних витрат, однак, для кожного з них характерна неоднозначність розмітки стартових пікселів в реальних ситуаціях через необхідність застосування евристик (вибір порогів для збігу яскравості, вибір цифрових масок і т.і.). Заслужовує на увагу у зв'язку з цим запропонований метод багатозначною розмітки, заснований на комбінації різних прийомів для зниження невизначеності. Важливе практичне значення мають алгоритми, що допускають паралельну обробку та прискорення процесу розмітки на основі логічного аналізу сусідніх елементів.

Оскільки сегментація зазвичай використовується не самостійно, а як частину певної системи (наприклад, системи машинного зору), то з

практичної точки зору, якість роботи методу оцінюється виходячи з роботи системи в цілому. Тому один і той же метод сегментації може виявитися добрим для однієї задачі і поганим для іншої. Для грубої оцінки якості методу в конкретній задачі зазвичай фіксують кілька властивостей, якими повинна володіти хороша сегментація. Якість роботи методу оцінюється в залежності від того, наскільки отримана сегментація володіє цими властивостями. Найбільш часто використовуються такі властивості: однорідність регіонів (однорідність кольору або текстури), несхожість сусідніх регіонів, гладкість кордону регіону, маленька кількість дрібних «дірок» всередині регіону. Різні методи сегментації орієнтовані на різні властивості розбиття. Тому при виборі методу сегментації для вирішення конкретного завдання, слід визначитися, які властивості розбиття дійсно важливі. У деяких прикладних задачах достатньо того, щоб розбиття характеризувалося лише першими двома з перелічених властивостей.

У системах технічного зору, складовою частиною яких є комп'ютерний зір, сегментація - це процес поділу цифрового зображення на декілька сегментів (множини пікселів, також званих суперпікселів). Мета сегментації полягає у спрощенні або зміні подання зображення, щоб його було простіше і легше аналізувати. Сегментацією називається процес поділу сцени на складові частини або об'єкти. Алгоритми сегментації, як правило, ґрунтуються на двох фундаментальних принципах: розривності і подібності. У першому випадку основний підхід ґрунтується на визначенні контурів, а в другому - на визначенні порогового рівня та розширенні області. Ці поняття застосовні як до статичних, так і до динамічних (залежних від часу) сцен.

Для зображень сегментацію найбільш часто проводять по яскравості для одноколірного зображення і колірним координатам для кольорового зображення. Розглянемо методи, що використовуються для проведення сегментації різних типів зображень.

Сегментація зображень зазвичай використовується для того, щоб виділити об'єкти і межі (лінії, криві, і т. і.) на зображеннях. Більш точно,

сегментація зображень - це процес присвоєння таких міток кожного пікселя зображення, що пікселі з однаковими мітками мають спільні візуальні характеристики.

Результатом сегментації зображення є множина сегментів, які разом покривають все зображення, або множина контурів, виділених з зображення. Всі пікселі в сегменті схожі за деякою характеристикою або обчислений властивості, наприклад за кольором, яскравості або текстурі. Сусідні сегменти значно відрізняються по цій характеристиці.

Сучасний період розвитку економіки промислово розвинених країн характеризується переходом до другого етапу науково-технічної революції - науково-технологічного, на основі комп'ютеризації й інформатизації всього суспільного виробництва. Для реалізації цього етапу в усьому світі прийняті широкі програми розробки комп'ютерів п'ятого й роботів третього покоління. Останні являють собою адаптивні системи, гнучко програмовані, забезпечені засобами відчуття для одержання інформації про навколишнє середовище й предмети виробництва. Подальший розвиток засобів автоматизації пов'язаний із широким застосуванням роботів і робототехнічних комплексів, значне місце для яких займають системи технічного зору (СТЗ).

Сфера застосування СТЗ включає три основних аспекти: візуальний контроль; керування й регулювання; виміри й визначення положення.

Візуальний контроль вирішує найважливіше завдання перевірки якості продукції автоматичного виробництва. Завдання керування виникають часто в процесі збору або сортування, складання або добування зі складу. У цьому випадку необхідно вміти розглянути й обійти перешкоди, не зачепити при переносі інші деталі й т.д. Це завдання є одним із найважчих при реалізації СТЗ. Основними є такі функції СТЗ: одержання зображення в полі зору; визначення наявності необхідних об'єктів; розпізнавання й виділення заданого об'єкта на зображенні; визначення координат об'єкта або його характерних точок щодо координатної системи датчика зображення; формування сигналів керування.

Призначення пристрою первинної обробки заключається в зменшенні загального часу обробки зображення. При цьому в більшості випадків здійснюється фіксація й подання зображення у вигляді характерних точок. Таким чином, на першому етапі здійснюється безпосередня обробка й подання інформації у вигляді, зручному для подальших стандартних перетворень, які звичайно виконують на універсальній ЕОМ. Сама ж попередня обробка враховує специфіку об'єкта, тому пристрій первинної обробки, як правило, є спеціалізованим, що реалізує необхідні алгоритми. Наявність у сучасних СТЗ пристроїв первинної обробки обумовлено вимогами високої швидкості обробки даних, технологічній гнучкості й високій роздільній здатності. Відсутність цього пристрою в системі збільшує загальний час обробки.

Таким чином, перш ніж піддатися аналізу, зображення повинне пройти етап підготовки, що складається з виконання операцій поліпшення візуальної якості (підвищення контрасту, усунення розмитості, підкреслення границь, фільтрація) і операцій формування графічного препарату (сегментація, виділення контурів) зображення. Всі ці етапи використовуються при обробці від початкового етапу сприйняття поля зору за допомогою датчиків, наприклад, телекамери, до кінцевого, котрим є розпізнавання зображення.

З метою класифікації методів і підходів, що використовуються в системах технічного зору, зір розбито на три основних підкласу: зір низького, середнього та високого рівнів. Системи технічного зору низького рівня призначені для обробки інформації з датчиків відчуттів.

Системи технічного зору середнього рівня пов'язані із завданнями сегментації, опису та розпізнавання окремих об'єктів. Ці завдання охоплюють множину підходів, заснованого на аналітичних представленнях об'єктів сцени, що аналізується.

Системи технічного зору високого рівня можна віднести до класу «інтелектуальних» машин і вони мають такі ознаки (признаками інтелектуальної поведінки): а) можливістю виділення істотної інформації з

множини незалежних ознак; б) здатністю до навчання на прикладах і узагальнення цих знань з метою їх застосування в нових ситуаціях; в) можливістю відновлення подій за неповною інформацією; г) здатністю визначати цілі і формулювати плани для досягнення цих цілей.

Перш за все необхідно подолати складнощі, пов'язані з сегментацією зображення. Будь-яке зображення, як правило, містить множину об'єктів, тому необхідно спочатку розбити його на підмножини пікселів, відповідних окремих об'єктах. А після розбиття зображення на ділянки можна ввести дані про цих ділянках або сукупностях ділянок в класифікатор для визначення міток об'єктів. На жаль, процес сегментації "знизу вгору" чреватий помилками, тому в якості альтернативного підходу може бути передбачений пошук для визначення груп об'єктів "зверху вниз". Це означає, що можна проводити пошук підмножини пікселів, які можна класифікувати як особа, і в разі успішного виконання даного етапу результатом стає успішне виявлення групи. Але підходи, засновані виключно на пошуку "зверху вниз" (або низхідному пошуку), мають високу обчислювальну складність, оскільки в них необхідно досліджувати вікна зображення різних розмірів, що знаходяться в різних місцях, а також порівнювати їх все з даними різних гіпотез про наявність об'єктів. В даний час така спадна стратегія використовується в більшості практично застосовуваних систем розпізнавання об'єктів, але подібна ситуація може змінитися в результаті удосконалення методів пошуку "знизу вгору" (висхідного пошуку).

В якості першоджерел інформації для дослідження характерних ознак технологічних об'єктів використовуються, як правило, відео- та фото зображення. При цьому на них можуть бути присутні об'єкти довільного виду. Тому перед безпосереднім аналізом зображених предметів необхідно виконати ряд попередніх операцій, що дозволяють отримати зображення самих об'єктів без сторонніх зображень.

У більшості автоматизованих систем обробки зображень виконується розпізнавання окремих об'єктів (їх фрагментів) по заданих об'єктно-

геометричним параметрам. При цьому зображення, яке обробляється локально неоднорідне і розпізнаваний об'єкт може ділитися на підоб'єкти і елементарні об'єкти.

Для ефективного використання таких зображень використовуються різні підходи до декомпозиції моделі даних, що дозволяють представити загальну модель як сукупність ієрархічно взаємозалежних простіших моделей різного рівня ієрархії.

Одним з найпоширеніших методів вирішення цих завдань є контурна сегментація. Методи контурної сегментації використовуються в багатьох областях, де об'єкти на аналізованих зображеннях володіють великою складністю і багатофакторністю, що обумовлює високі вимоги до надійності, точності та достовірності результатів досліджень. Використання обчислювальної техніки та математичних методів в цій області дозволяє не тільки прискорити процес обробки матеріалу, а й підвищити точність результатів дослідження.

Швидкий розвиток цифрової техніки останнім часом відкриває нові можливості в реалізації цих методів. Збільшення швидкодії обчислювальної техніки дозволяє використовувати складні, критичні до часу алгоритми, а завдяки появі кольорових телевізійних датчиків високого дозволу можна отримувати і обробляти кольорові зображення. Саме нові технічні можливості дозволяють значно розширити коло досліджень, відкривають нові шляхи вирішення завдань, що стосуються аналізу зображень.

Найбільш типові методи контрастної сегментації: метод певного вікна і метод ланцюгових кодів.

Кордон - контрастна область зображення, що містить різку відмінність яскравості між двома сусідніми пікселями. Такі перепади яскравості, як правило, є межами об'єкту, де фон і яскравість самого об'єкта значно відрізняються. Існує множина різних методів виділення меж. Вони можуть поєднуватися з корекцією по гістограмах і бінаризації зображення. Найпоширеніші методи виділення меж: алгоритми Собеля, Канні, Робертса

(рисунок 1.1).



Рисунок 1.1 – Маски поширених операторів

На підготовчому етапі методу певного вікна знаходять області, що містять необхідну контрастність (високу або, навпаки, низьку). Далі створюється вікно, виходячи з примірних розмірів і форми шуканого об'єкта, і вважається кількість граней в «підозрілих» областях. Якщо воно перебуває в заданому діапазоні - об'єкт виділений. Діапазон кількості граней вибирається експериментально.

Недолік використання даного методу полягає в тому, що обраний поріг бінаризації не може працювати добре на всіх зображеннях. Освітлення, кольори можуть впливати на якість бінаризації зображення. Ще один недолік методу - низька швидкість роботи, так як значення всіх пікселів у виділеному

вікні мають бути неодноразово підсумовані. Третій істотний недолік - низька якість розпізнавання. Тому необхідно додатково використовувати інші методи.

Метод контурної сегментації здатний визначити місцезнаходження об'єкта з вірогідністю близько 73%.

Метод ланцюгового кодування був запропонований Фріманом. Він полягає в тому, щоб кордон об'єкта, розташованого на дискретній сітці, представити у вигляді набору елементарних відрізків. Тоді повною характеристикою межі об'єкта в кожній точці є напрямок необхідного відрізка (1 - вгору, 2 - вправо, 3 - вниз, 4 - вліво). Передбачається, що точки на кордоні є тільки 4-х зв'язаними (модифікація методу використовує 8-зв'язну модель).

Безсумнівним достоїнством подання кордону зображуваного об'єкта ланцюговим кодом є простота реалізації алгоритму його опису, простота отримання на основі цього опису геометричних характеристик об'єкту (периметр, площа, лінійні розміри по вертикалі і горизонталі та інше), можливість досягнення інваріантності до перетворень подібності - масштабування зображення, його переносу і повороту.

Основним недоліком даного методу є висока нестійкість отримуваних описів до спотворень в зображеннях.

1.2 Сегментація зображень та методи машинного навчання

Методи сегментації можна розділити на два класи: автоматичні - які не потребують взаємодії з користувачем і інтерактивні - які використовують для користувача введення безпосередньо в процесі роботи. В даній роботі будемо розглядати тільки автоматичні методи.

Задачі автоматичної сегментації діляться на два класи:

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

Між цими двома постановками завдання є принципова різниця. У першому випадку завдання сегментації полягає у пошуку певних областей, про які є апіорна інформація (наприклад, ми знаємо колір, форму областей, або цікаві нас області представляють собою зображення відомого об'єкта). Методи цієї групи вузько спеціалізовані для кожної конкретної задачі. Сегментація в такій постановці використовується в основному в задачах машинного зору (аналіз сцен, пошук об'єктів на зображенні).

У другому випадку ніяка апіорна інформація про властивості областей не використовується, зате на саме розбиття зображення накладаються деякі умови (наприклад, всі області повинні бути однорідні за кольором і текстурою). Так як при такій постановці задачі сегментації не використовується апіорна інформація про зображених об'єктах, то методи цієї групи універсальні і застосовні до будь-яких зображень. В основному сегментація в цій постановці застосовується на початковому етапі рішення задачі, для того щоб отримати уявлення зображення в більш зручному вигляді для подальшої роботи.

Складовою частиною задач обробки та аналізу зображень є сегментація, тобто поділ зображення на області, для яких виконується певний критерій однорідності, наприклад, виділення на зображенні областей приблизно однаковою яскравості. Поняття області зображення використовується для визначення зв'язної групи елементів зображення, що мають певну загальну ознаку (властивість).

Метою сегментації є розбиття образу на області. Розглянемо методи сегментації, засновані на прямому знаходженні областей.

Нехай S - область образу.

Розглянемо сегментацію як процес розбиття образу S на n під областей S_1, S_2, \dots, S_n , так що

$$S = \{S_i\}, i = \overline{1, N} \quad (1.1)$$

S_i -зв'язана область, $i = 1, 2, \dots, n$,

$$I = \bigcup_{i=1..N} S_i \quad (1.2)$$

де I – стосовно зображення параметр, який описує яскравість або інший параметр зображення,

Для всіх i та j виконується умова

$$\forall i, j = \overline{1, N} : i \neq j S_i \cap S_j = \emptyset \quad (1.3)$$

$$\forall i = \overline{1, N}, P(S_i) = \text{істина} \quad (1.4)$$

де $P(S_i)$ - логічний предикат, визначений на точках з множини S_i .

$$\forall i, j = \overline{1, N} : i \neq j P(S_i \cup S_j) = \text{невірно} \quad (1.5)$$

Умова 1 означає, що сегментація повинна бути повною, тобто кожен піксель має перебувати в образі. Друга умова вимагає, щоб точки в області були зв'язаними. Умова 3 вказує на те, що області не повинні перетинатися. Умова 4 визначає властивості, яким повинні задовольняти пікселі в сегментованій області. Простий приклад: $P(S_i) = \text{ІСТИНА}$, якщо всі пікселі в S_i мають однакову інтенсивність. Умова 5 означає, що області S_i і S_j розрізняються по предикату P .

Сегментація зображення являє собою поділ або розбиття зображення на області за подібністю властивостей їх точок. Найбільш часто сегментацію проводять по яскравості для одноколірного зображення і колірним координатам для кольорового зображення. Застосовується також сегментація, заснована на контурах, сегментація, при якій в якості розділового ознаки використовується текстура і сегментація за формою.

На теперішній час існують багато різних способів сегментації зображень. Їх можна розділити на такі групи:

- порогові методи;
- методи нарощування областей (методи водорозділу, розщеплення - злиття, центроїдного зв'язування);
- текстурні методи;
- методи виділення границь;
- методи кластерного аналізу;
- графові методи.

Багато зображень можна охарактеризувати тим, що вони містять певний цікавий для нас об'єкт досить однорідною яскравості на тлі іншої яскравості. Типовими прикладами можуть служити машинописні і рукописні тексти, медичні проби під мікроскопом, літаки на злітній смузі. Для таких зображень яскравість служить відмітною ознакою, який можна використовувати для локалізації об'єкта. Якщо нас цікавить об'єкт має білий колір і розташований на чорному тлі або навпаки, то визначення точок об'єкта є тривіальну задачу встановлення порогу по середній яскравості. На практиці, проте, зустрічаються певні труднощі, наприклад, коли спостерігається зображення піддається впливу шуму, причому як на об'єкті, так і на тлі допускається деякий розкид значень яскравості. Інша часто зустрічається складність полягає в тому, що фон може бути неоднорідним.

Спочатку розглянемо постановку задачі сегментації, як розбиття зображення на однорідні області. Така постановка виникла раніше, ніж завдання виділення областей зображення з відомими властивостями, і методи цієї групи на даний момент добре розроблені.

Відомо кілька аналітичних підходів до порогового обмеження яскравості. Один з методів полягає у встановленні порогу на такому рівні, при якому загальна сума елементів з підпороговою яскравістю узгоджена з апіорними ймовірностями цих значень яскравості. Наприклад, може бути відомо, що чорні знаки займають 25% загальної площі машинописного

сторінки. Тоді пороговий рівень для такого зображення необхідно встановити таким, щоб яскравість чверті елементів була нижче порога. Інший підхід до порогового обмеження по яскравості полягає у виборі порога, відповідного мінімуму біомодальної гістограми, що знаходиться між двома її піками. Визначення цього мінімального значення часто утруднене внаслідок ступінчастості гістограми. Тому ділянку гістограми між піками апроксимується деякою аналітичною функцією і знаходиться її мінімум шляхом обчислення її похідних.

Розглянемо область зображення в районі об'єкта, де яскравість збільшується з рівня низького "плато" до рівня більш високого "плато", з'єднаних похилою поверхнею. На плоских ділянках лапласіан дорівнює нулю, а вздовж похилої поверхні - майже нулю. В області переходу від низького "плато" лапласіан буде мати велике позитивне значення, а при переході до високого "плато" - велике негативне значення. Гістограма, побудована з використанням лише точок вихідного зображення, які відповідають дуже високим або дуже низьким значенням лапласіана, виявляється бімодальною з виразною "долиною" між піками. Визначенню порога яскравості можуть допомогти і інші процедури виділення перепадів.

Якщо фон на зображенні неоднорідний, то необхідно зробити так, щоб поріг по яскравості адаптувався до її середнього рівня. Це можна виконати, розбиваючи зображення на маленькі фрагменти і встановлюючи свої пороги для кожного фрагмента. Пороги для кожного елемента зображення можна потім визначити шляхом інтерполяції між центрами фрагментів.

Ефективну сегментацію можна виконати для деяких класів зображень за допомогою рекурентного методу порогового обмеження по яскравості. На першому етапі виконується порогове обмеження зображення шляхом знаходження мінімального значення між модами гістограми яскравості для того, щоб відокремити більш яскраві області від більш темних. Потім формуються гістограми для кожної сегментованої частини. Якщо ці гістограми неунімодальні, то отримані сегменти можна знову піддати

пороговому обмеженню. Процес продовжується до тих пір, поки гістограми для всіх сегментів не стануть унімодальному або ж не буде досягнутий певний бажаний рівень сегментації.

Методи з використанням гістограм дуже ефективні, коли порівнюються з іншими методами сегментації зображень, тому що вони вимагають тільки один прохід по пікселях. У цьому методі гістограма обчислюється за всіма пікселям зображення і її мінімуми і максимуми використовуються, щоб знайти кластери на зображенні. Колір або яскравість можуть бути використані при порівнянні.

Поліпшення цього методу - рекурсивно застосовувати його до кластерів на зображенні для того, щоб поділити їх на більш дрібні кластери. Процес повторюється з усе меншими й меншими кластерами до тих пір, коли перестануть з'являтися нові кластери.

Один недолік цього методу - те, що йому може бути важко знайти значні мінімуми і максимуми на зображенні. У цьому методі класифікації зображень схожі метрика відстаней і зіставлення інтегрованих регіонів.

Підходи, засновані на використанні гістограм, можна також швидко адаптувати для декількох кадрів, зберігаючи їх перевагу в швидкості за рахунок одного проходу. Гістограма може бути побудована кількома способами, коли розглядаються декілька кадрів. Той же підхід, який використовується для одного кадру, може бути застосований для кількох, і після того, як результати об'єднані, мінімуми і максимуми, які було складно виділити, стають більш помітними. Гістограма також може бути застосована для кожного пікселя, де інформація використовується для визначення найбільш частого кольору для даного положення пікселя. Цей підхід використовує сегментацію, засновану на рухомих об'єктах і нерухомому оточенні, що дає інший вид сегментації, корисний у відеотрекінгу.

Оскільки сегментація зазвичай використовується не самотійно, а як частину певної системи (наприклад, системи машинного зору), то з практичної точки зору, якість роботи методу оцінюється виходячи з роботи

системи в цілому. Тому один і той же метод сегментації може виявитися добрим для однієї задачі і поганим для іншої.

Для грубої оцінки якості методу в конкретній задачі зазвичай фіксують кілька властивостей, якими повинна володіти хороша сегментація. Якість роботи методу оцінюється в залежності від того, наскільки отримана сегментація володіє цими властивостями. Найбільш часто використовуються такі властивості: однорідність регіонів (однорідність кольору або текстури), несхожість сусідніх регіонів, гладкість кордону регіону, маленька кількість дрібних «дірок» всередині регіону і т. д.

Різні методи сегментації орієнтовані на різні властивості розбиття. Тому при виборі методу сегментації для вирішення конкретного завдання, слід визначитися, які властивості розбиття дійсно важливі.

1.3 Методи сегментації границь зображень

Досить часто при аналізі зображень виникає завдання поділу пікселів зображень на групи за деякими ознаками. Такий процес розбиття на групи називається сегментацією. Найбільш відомими є два види сегментації - сегментація по яскравості для бінарних зображень і сегментація по колірним координатам для кольорових зображень. Методи сегментації можна розглядати як формалізацію поняття виділення об'єкта з фону або понять пов'язаних з градієнтом яскравості. Алгоритми сегментації характеризуються деякими параметрами надійності та достовірності обробки.

Зазвичай аналіз зображень включає такі операції, як отримання зовнішнього контуру зображених об'єктів і запис координат точок цього контуру. Часто потрібно отримати зовнішній контур у вигляді замкнутої кривої або сукупності відрізків дуг. Є три загальних підходи до подання меж об'єкту: апроксимація кривих, простежування контурів і зв'язування точок перепадів.

Метод апроксимації кривих можна використовувати для контурної

сегментації шляхом підбору аналітично заданої кривої до сукупності точок контурного препарату, отриманого за допомогою детектора перепадів. Якщо є апіорна інформація щодо очікуваної форми об'єкта, наприклад, відомо, що це прямокутник або коло, то в якості апроксимуючої кривої потрібно взяти прямокутник або коло. Або ж можна спробувати застосувати апроксимацію поліномами, використовуючи ітеративні методи підбору прямих ліній. Метод підбору кривих досить добре підходить для об'єктів із простою структурою. Труднощі виникають, коли зображення містить множина об'єктів, що перекриваються, і точки перепадів утворюють гіллясті структури.

Простежування контурів нагадує поведінку жука, що обходить перешкоди. У випадку двійкового зображення, уявний жук починає свій шлях на білому полі і рухається у напрямку до області чорних елементів зображення. Після того, як жук перетне чорний елемент, він повертається ліворуч і переходить до наступного запису. Якщо цей елемент чорний, жук знову повертається ліворуч, якщо ж елемент виявляється білим, то жук повертається направо. Ця процедура продовжується до тих пір, поки жук не повернеться у вихідну точку. Декартові координати точок переходу з чорного на біле або з білого на чорне дають розташування кордону.

Інші труднощі пов'язані з наявністю "дірок" всередині об'єкта, які жук може пропустити. Проблеми такого характеру можна подолати, забезпечивши жука деякої пам'яттю та інтелектом, що дозволить йому запам'ятовувати свої минулі кроки і відновлювати простежування, якщо даний шлях виявляється помилковим. Описаний алгоритм найчастіше застосовується до зображень, підданих пороговому обмеженню, але дану ідею легко можна поширити і на багато градаційне зображення. Один з підходів полягає в тому, що гранична точка відзначається, якщо різниця яскравості сусідніх елементів зображення досить велика. У результаті жук формує контури по ходу свого руху.

Існує багато зображень, які містять досліджуваний об'єкт досить однорідною яскравості на тлі іншої яскравості. Як приклад можна навести

рукописний текст, ряд медичних зображень і т.д. Якщо яскравості точок об'єкта різко відрізняються від яскравості точок фону, то рішення задачі встановлення порогу є нескладним завданням. На практиці це не так просто, оскільки досліджуване зображення піддається впливу шуму і на ньому допускається деякий розкид значень яскравості. Відомо кілька аналітичних підходів до порогового обмеження по яскравості. Один з методів полягає у встановленні порогу на такому рівні, при якому загальна сума елементів з підпороговою яскравістю узгоджена з апіорними ймовірностями цих значень яскравості.

Аналогічні підходи можна застосувати для обробки кольорових і спектрозональних зображень. Існує також такий вид сегментації як контурна сегментація. Досить часто аналіз зображень включає такі операції, як отримання зовнішнього контуру зображень об'єктів і запис координат точок цього контуру. Відомо три основних підходи до подання меж об'єкту: апроксимація кривих, простежування контурів і зв'язування точок перепадів. Найбільш простим видом сегментації є порогова сегментація. Вона знайшла дуже широке застосування у робототехніці. Це пояснюється тим, що в цій сфері зображення досліджуваних об'єктів, у своїй більшості, мають досить однорідну структуру і різко виділяються їх фону. Але крім цього, для достовірної обробки потрібно знати, що зображення складається з одного об'єкту і фону, яскравості яких знаходяться в строго відомих діапазонах і не перетинаються між собою.

Об'єкт або область зображення, яку бажано виділити, звичайно відокремлюється від сусідніх частин зображення перепадом яскравості. Тому з'єднання точок перепадів можна використовувати для сегментації. Метод з'єднання точок перепадів, розроблений Робертсом, заснований на тих же принципах, що і більшість інших методів такого зв'язування.

У методі Робертса аналогові значення градієнтів (різниці яскравості сусідніх елементів) аналізуються блоками розміром 4x4 елемента. Елемент з найбільшим у блоці значенням модуля градієнта вважається пробної

контурній точкою, якщо це значення більше порога. Потім до значень градієнта близько цієї пробної точки підбираються лінії довжиною в п'ять елементів з орієнтацією "північ", "схід", "південь", "захід". Якщо відношення найкращої апроксимації до найгіршої більше другого порога, то пробна контурна точка оголошується дійсною і йому приписується орієнтація, відповідна найкращої апроксимації. Далі до пар контурних точок підбираються прямі лінії, якщо ці точки розташовані у суміжних блоках розміром 4x4 елемента і якщо напрямок кожної лінії перебуває в діапазоні $\pm 23^\circ$ відносно орієнтації контуру в кожній контурній точці. Точки, що не задовольняють критеріям з'єднання, відкидаються. Маленькі трикутники виключаються викреслюванням найбільшої сторони, а маленькі чотирикутники замінюються їх найбільшою діагоналлю. Короткі виступаючі лінії також знищуються. На цьому етапі короткі розриви замінюються з'єднаннями з прямих ліній. Цей вид з'єднання точок перепаду можна використовувати для широкого класу детекторів перепаду.

При цьому підході завдання сегментації формулюється як задача пошуку меж регіонів. Методи пошуку кордонів добре розроблені для напівтонових зображень. Півтонове зображення розглядається як функція двох змінних (x і y), і передбачається, що межі регіонів відповідають максимумів градієнта цієї функції. Для їх пошуку застосовується апарат диференціальної геометрії (у найпростішому випадку це фільтри Roberts, Kirsch, Prewitt, Sobel).

Для підвищення стійкості до шуму, перед застосуванням фільтрації зображення звичайно розмивають. Завдяки комутативності оператора Лапласа і Гауссова фільтра, можна одночасно здійснювати розмиття і пошук кордонів. У методі Canny комбінуються результати пошуку кордонів при різному ступені розмиття.

Інший підхід заснований на застосуванні steerable filters, які здійснюють диференціювання за напрямком. Для таких фільтрів можна вибрати базис, через який виражається диференціювання за будь-якого

напрямку. Для пошуку кордонів комбінуються результати застосування базисних фільтрів. У [20] пропонується комбінований підхід, що використовує steerable filters і cellular neural networks.

Розглянемо деякі методи виявлення трьох основних видів розривів яскравості, що зустрічаються в цифрових зображеннях: точок, ліній і перепадів. Найбільш загальним способом пошуку розривів є обробка зображення за допомогою ковзної маски. Для показаної на рисунку 1.1 маски розмірами 3x3 елементи ця процедура заснована на обчисленні лінійної комбінації коефіцієнтів маски зі значеннями яскравості елементів зображення, що покриваються маскою. Інакше кажучи, при використанні цієї маски відгук у кожній точці зображення задається виразом

$$R = w_1 z_1 + w_2 z_2 + \dots + w_9 z_9 = \sum_{i=1}^9 w_i z_i \quad , \quad (1.6)$$

де z_i -...значення яскравості пікселя, що відповідає коефіцієнтові w_i маски. Як звичайно, відгук маски приписується позиції її центрального елемента. У принципі, виявлення окремих ізольованих точок на зображенні не представляє складності. Скористаємося маскою, показаної на рисунку 1.2, і будемо вважати, що в тому пікселі, куди попадає центр маски, виявлена точка, якщо

$$|R| \geq T \quad (1.7)$$

де T – додатній поріг, а R обчислюється відповідно до (1.6). По суті, у даній формулі виміряється зважена сума різниць значень центрального елемента і його сусідів. Ідея полягає в тому, що ізольована точка, що розташована в однорідній або майже однорідної області точка, значення яскравості якої істотно відрізняється від навколишнього тла буде помітно відрізнятися по яскравості від найближчих сусідів, а виходить, буде легко виявлятися за допомогою маски приведенного виду.

Однак у даному випадку ціль складається винятково у виявленні

окремих точок, тому інтерес представляють тільки досить великі розходження (обумовлені порогом T), при яких точка може вважатися ізольованою. Звернемо увагу, що сума коефіцієнтів маски дорівнює нулеві, так що на областях постійної яскравості вона буде давати нульовий відгук.

w_1	w_2	w_3
w_4	w_5	w_6
w_7	w_8	w_9

Рисунок 1.2 – Загальне представлення маски розмірами 3x3

-1	-1	-1
-1	8	-1
-1	-1	-1

Рисунок 1.3 – Маска для виявлення точок

Даний спосіб виявлення ізольованих точок досить специфічний, оскільки передбачається одноелементний розрив, навколо якого (в області маски детектора) розташовується однорідне тло. Якщо така умова не дотримується, то для виявлення розривів яскравості більш підходящими будуть інші методи, розглянуті далі в цьому розділі.

Наступної за рівнем складності є виявлення ліній. Розглянемо набір масок, показаний на рис. 1.4. При ковзанні першої маски по зображенню, найбільш сильний відгук буде на горизонтальних лініях товщиною в один піксель; причому, якщо яскравість тла однакова, то відгук буде максимальним, коли лінія проходить горизонтально через центр маски. Аналогічні експерименти підтвердять, що друга маска на рис. 1.4 дає найбільший відгук на лініях, що проходять під кутом $+45^\circ$; третя - на вертикальних лініях; четверта – по лініях під кутом -45° . Ці напрямки можна

виявити і по тій ознаці, що кращі напрямки кожної з масок характеризуються великими значеннями вагових коефіцієнтів (а саме, 2), чим будь-які інші напрямки.

Позначимо через R_1, R_2, R_3 і R_4 відгуки масок, показаних на рис. 1.4 (зліва направо), де значення R_i обчислюються відповідно до співвідношення (1.6). Будемо вважати, що зображення обробляється незалежно за допомогою кожної з цих масок. Якщо в деякій точці зображення $|R_i| > |R_j|$ для усіх $j \neq i$ то ця точка, швидше за все, зв'язана з лінією, орієнтованої уздовж напрямку маски. Наприклад, якщо в якійсь точці зображення $|R_i| > |R_j|$ для $j = 2, 3, 4$, то ця точка, очевидно, належить горизонтальній лінії. Альтернативна задача може бути сформульована у виді пошуку ліній, що йдуть у заданому напрямку. У такому випадку можна обробити все зображення за допомогою маски для цього напрямку, застосовуючи граничне перетворення (1.6) до одержуваного відгуку. Іншими словами, якщо інтерес представляють усі лінії в зображенні, котрі орієнтовані по напрямку даної маски, досить пройти цією маскою по всьому зображенню, порівнюючи абсолютне значення результату з заданим порогом. Точки, що залишилися при цьому, відповідають найбільшим значенням відгуку, що у випадку ліній товщиною в один піксель найбільш близькі до напрямку, обумовленому маскою.

-1	-1	-1	-1	-1	2	-1	2	-1	2	-1	-1
2	2	2	-1	2	-1	-1	2	-1	-1	2	-1
-1	-1	-1	2	-1	-1	-1	2	-1	-1	-1	2

Рисунок 1.4 – Маски для виявлення ліній

Хоча виявлення точок і ліній відіграє важливу роль у будь-якому розгляді задачі сегментації зображень, виявлення перепадів яскравості є набагато більш загальним підходом до перебування інтерпретованих

розривів на яркій картині.

Ідеальний перепад має властивості моделі у вигляді ступінчатого стрибка. Відповідно до цієї моделі, ідеальний контурний перепад - це множина з'єднаних пікселів (у даному випадку по вертикалі), кожний з яких розташований поруч із прямокутним стрибком яскравості.

На практиці оптичні обмеження, дискретизація, а також недосконалість інших елементів системи реєстрації зображень приводять до одержання розмитих перепадів яскравості. Причому ступінь розфокусування визначається такими факторами, як якість системи реєстрації, крок дискретизації й умови висвітлення, при яких зображення було отримано. У результаті перепади яскравості більш точно моделюються похилим профілем. Крутість похилої ділянки назад пропорційна ступені фокусування перепаду. У такій моделі вже більше немає тонкої траєкторії (шириною в один піксель). Замість цього точкою перепаду яскравості тепер є будь-як точка, що лежить на похилій ділянці профілю, а сам перепад являє собою зв'язну множину, утворена всіма такими точками. "Ширина" такого перепаду визначається довжиною похилої ділянки, на якому здійснюється перехід від початкової яскравості до кінцевого. Ця довжина залежить від крутості ділянки, що, у свою чергу, визначається ступенем розфокусування перепаду. Така залежність виглядає цілком осмислено: розмиті перепади виглядають широкими, а різкі перепади - тонкими.

Друга похідна позитивна в точці переходу від темної ділянки до похилого, негативна в точці переходу від похилої ділянки до світлого, і дорівнює нулеві на лінійному схилі і ділянках постійної яскравості. У випадку зворотного перепаду яскравості (від світлого до темного) знаки похідних зміняться на протилежні.

З проведеного розгляду можна укласти, що значення першої похідної може використовуватися для виявлення наявності перепаду яскравості в кожній точці зображення (тобто з'ясування, чи знаходиться точка на похилій ділянці). Аналогічно, знак другої похідної дозволяє визначити, чи лежить

піксель, що знаходиться на перепаді, на темній або світлій його частині.

Звернемо увагу на дві додаткових властивості другої похідної поблизу перепаду яскравості: (1) вона дає два ненульових (позитивні і негативні) значення для кожного перепаду, що є небажаною властивістю; і (2) уявлювана пряма лінія, що з'єднує максимальні позитивне і негативне значення другій похідній поблизу перепаду, перетинає нульовий рівень приблизно в середині перепаду яскравості. Це властивість перетинання нульового рівня другій похідній досить корисно для локалізації середини широких перепадів.

Як цього можна було очікувати, друга похідна виявляється ще більш чутливою до шуму. Таким чином, важливо пам'ятати, що навіть невеликий шум може впливати на першу і другу похідні, застосовувані для виявлення перепадів на зображеннях. Зокрема, у практичних задачах, де можлива поява помітного шуму, доцільно розглянути питання про згладжування зображення перед обчисленням похідних.

Похідні першого порядку в зображенні обчислюються за допомогою градієнта. Для одержання похідних другого порядку застосовується лапласіан. Ясно, що вибір критерію для граничних точок визначає якість роботи крайової методів. В для вибору оптимального критерію граничних точок застосовується машинне навчання.

Основною проблемою методів пошуку кордонів є нестійкість до шуму. Крім того, оскільки поняття кордону своє для кожної задачі, кожен раз при застосуванні методів пошуку кордонів потрібно додатково вибирати метод доопрацювання результатів фільтрації (edge linking).

1.4 Методи нарощування областей

Методи цієї групи враховують просторове розташування точок в певному напрямку.

Методи вирощування регіонів засновані на наступній ідеї. Спочатку по

деякому правилу вибираються центри регіонів (seeds), до яких поетапно приєднуються сусідні точки, що задовольняють деякому критерію. Процес вирощування регіонів (region growing) зупиняється, коли ні одна точка зображення не може бути приєднана до жодного регіону.

Застосовуються різні критерії, на підставі яких точка приєднується або не приєднується до регіону: близькість (в певному сенсі) точки до центру регіону; близькість до сусідньої точки, приєднаної до регіону на попередньому кроці; близькість за деякою статистикою регіону; вартість найкоротшого шляху від точки до центру регіону, і т. п. Існують різні стратегії вибору зерен (seeds) і вирощування регіонів.

Методи дроблення-злиття складаються з двох основних етапів: дроблення і злиття. Дроблення починається з деякого розбиття зображення, не обов'язково на однорідні області. Процес дроблення областей відбувається до тих пір, поки не буде отримано розбиття зображення (пересегментація), що задовольнить властивості однорідності сегментів. Потім відбувається об'єднання схожих сусідніх сегментів до тих пір, поки не буде отримано розбиття зображення на однорідні області максимального розміру.

Ще один метод – метод керованого вододілу. Пропонується розглядати зображення як деяку карту місцевості, де значення яскравості являють собою значення висот щодо деякого рівня. Якщо цю місцевість заповнювати водою, тоді утворюються басейни. При подальшому заповненні водою, ці басейни об'єднуються. Місця об'єднання цих басейнів відзначаються як лінії вододілу.

Поділ дотичних предметів на зображенні є однією з важливих задач обробки зображень. Часто для вирішення цього завдання використовується так званий метод маркерного вододілу. При перетвореннях з допомогою цього методу потрібно визначити "водозбірні басейни" і "лінії вододілу" на зображенні шляхом обробки локальних областей в залежності від їх характеристик яскравості.

Метод маркерного вододілу є одним з найбільш ефективних методів сегментації зображень. При реалізації цього методу виконуються наступні основні процедури:

Обчислюється функція сегментації. Вона стосується зображень, де об'єкти розміщені в темних областях і є важко помітними.

Обчислення маркерів переднього плану зображень. Вони обчислюються на підставі аналізу зв'язності пікселів кожного об'єкта.

Обчислення фонових маркерів. Вони являють собою пікселі, які не є частинами об'єктів.

Модифікація функції сегментації на підставі значень розташування маркерів фону і маркерів переднього плану.

Обчислення на підставі модифікованої функції сегментації.

У даному прикладі серед функцій пакету Image Processing Toolbox найбільш часто використовуються функції `fspecial`, `imfilter`, `watershed`, `label2rgb`, `imopen`, `imclose`, `imreconstruct`, `imcomplement`, `imregionalmax`, `bwareaopen`, `graythresh` і `imimposemin`.

1.5 Текстура як цінна ознака якості зображення

Вже давно стало ясно, що текстура була би цінною ознакою для сегментації зображень. Проте застосування цієї ідеї на практиці перешкоджало досі відсутність надійних і ефективних способів виявлення та вимірювання параметрів текстури.

Один з підходів до текстурної сегментації полягає в підрахунку деякої міри зернистості текстури у всіх точках зображення з подальшим виявленням змін цього заходу. Фактично вхідне зображення піддається попередній обробці з тим, щоб перетворити його в зображення по яскравості. Основні труднощі, пов'язана з цим підходом, полягає в тому, що параметри текстури вимірюються в деякому вікні i , отже, їх значення, отримані в околиці кордону між текстурними областями, є усередненими. В результаті важко

точно локалізувати кордон між текстурними областями.

Інший підхід до текстурної сегментації полягає у виявленні перехідної зони між областями з різною текстурою. Основна ідея виявлення зміни текстури ідентична ідеї виявлення перепаду по яскравості: відмінність між текстурними областями посилюється у всіх точках зображення, а потім контрастували препарат піддається пороговому обмеження для локалізації текстурних перепадів. Параметри текстури обчислюються для кожного з чотирьох скануються по зображенню суміжних блоків розміром $W \times W$ елементів, і знаходиться сума модулів перехресних різниць, а потім виконується порогове обмеження для локалізації значних змін текстури. Цей метод можна узагальнити так, щоб виконувати обчислення по суміжних вікнам, об'єднаним у групи 3×3 . Тоді значення ознак текстури, обчислені в кожному вікні, можна об'єднати деяким лінійним або нелінійним способом, аналогічним методам контрастування перепадів яскравості по області 3×3 .

У задачах розпізнавання зображень часто виявляється корисним розбити об'єкт складної форми на зв'язний набір частин простий, легко описуваної форми. Наприклад, букву "Т" можна розбити на зв'язкові прямокутники. Сегментацію форми можна виконати в два етапи. Спочатку об'єкт довільної форми апроксимується набором прямолінійних або криволінійних відрізків. Потім апроксимувати форма піддається сегментації по точках перегину. Основні правила сегментації досить прості: найближчі точки угнутості з'єднують, утворюючи "перешийки", які потім видаляються з основного тіла; процес продовжується до тих пір, поки всі частини не виявляться опуклими.

1.6 Сегментація зображень на основі класичних методів

Методи теорії графів – один з методів, що найбільш активно розвиваються в напрямку сегментації зображень останнім часом. Загальна ідея методів цієї групи така.

Зображення представляється у вигляді зваженого графа, з вершинами у точках зображення. Вага ребра графа відображає подібність точок у деякому сенсі (відстань між точками за деякою метриці). Розбиття зображення моделюється розрізами графа.

Зазвичай в методах теорії графів вводиться функціонал «вартості» розтину, що відображає якість отриманої сегментації. Так завдання розбиття зображення на однорідні області зводиться до оптимізаційної задачі пошуку розрізу мінімальної вартості на графі. Такий підхід дозволяє крім однорідності кольору і текстури сегментів управляти також формою сегментів, їх розміром, складністю границь.

Для пошуку розрізу мінімальної вартості застосовуються різні методи: жадібні алгоритми (на кожному кроці вибирається таке ребро, щоб сумарна вартість розрізу була мінімальною), методи динамічного програмування (гарантується, що, вибираючи на кожному кроці оптимальне ребро, отримаємо в результаті оптимальний шлях), алгоритм Дейкстри, і т. п. Розглянемо деякі методи теорії графів докладніше. Метод Normalized Cut запропонований J. Shi, J. Malik. Оптимізація нормалізованого функціоналу зводиться до задачі пошуку власних значень матриці попарних відстаней між всіма точками зображення. Для сегментації зображення на дві частини досить знайти друге за величиною власне значення такої матриці. Складність ефективного алгоритму пошуку власних значень розрідженій матриці лінійна за кількістю точок зображення. Однак метод вимагає зберігання матриці розміром $n \times n$, де n - число точок зображення, і тому в початковому вигляді непридатний до великих зображень.

Для даного методу запропоновані модифікації, що дозволяють скоротити складність алгоритму і вимоги по пам'яті за рахунок апроксимації матриці відстаней. Такий підхід дає вигоду у швидкості роботи в 10-20 разів у порівнянні з вихідним методом.

Метод Nested Cuts запропонований Olga Veksler. Основний принцип цього методу полягає у відділенні кожної точки зображення від спеціальної

точки за межами зображення розрізом мінімальної вартості. При такому підході зображення ділиться на непересічні сегменти. Показано, що величиною сегментів зображення можна управляти, накладаючи обмеження на вартість розрізу. У статті описується ефективний алгоритм, заснований на властивостях розбиття. Однак цей метод працює вкрай повільно.

М. Pavan та М. Pelillo був запропонований новий підхід, заснований на розрізах графа. Автори вводять таке визначення сегменту, яке дозволяє переформулювати задачу пошуку розрізу на графі як завдання квадратичного програмування. Запропоновано метод розв'язання отриманої задачі, заснований на методах еволюційної теорії ігор. Цей підхід також вимагає зберігання в пам'яті матриці попарних відстаней, як і метод Normalized Cuts.

Метод сегментації SWA (Segmentation by Weighted Aggregation) заснований на групуванні схожих точок зображення. Основна ідея методу полягає в побудові піраміди зважених графів, кожен з яких отримано з попереднього шляхом об'єднання схожих вершин. На кожному кроці ваги зв'язків перераховуються. У процесі побудови піраміди обчислюються різні статистики, що характеризують форму, колір, текстуру регіонів, ці статистики використовуються для обчислення міри схожості регіонів. Потім, дотримуючись ідеології методів теорії графів, для отриманого графа вводиться функціонал вартості розрізу і шукається розріз мінімальної вартості. При цьому, на відміну від більшості методів теорії графів, SWA має складність $O(n)$, де n - число точок зображення, причому число операцій для кожної точки складає всього декілька десятків.

Методи розрізу графа можуть бути ефективно застосовані для сегментації зображень. В цих методах зображення представляється як виважений неорієнтований граф. Зазвичай, піксель або група пікселів асоціюється вершиною, а ваги ребер визначають (не) схожість сусідніх пікселів. Потім граф (зображення) розрізається згідно з критерієм, створеному для отримання «хороших» кластерів. Кожна частина вершин (пікселів), що отримується цими алгоритмами, вважається об'єктом на

зображенні. Деякі алгоритми цієї категорії - це нормалізовані розрізи графів, випадкове блукання, мінімальний розріз, ізопериметричний поділ і сегментація за допомогою мінімального кістяка.

Якість роботи методів теорії графів сильно залежить від вибору метрики. Тому для вибору оптимальної метрики в застосовують машинне навчання. Основні проблеми методів теорії графів - це низька швидкість роботи і великі витрати пам'яті. Більшість методів вимагає зберігання в пам'яті матриці попарних відстаней між точками зображення, розмір якої дорівнює квадрату числа точок. Такі обмеження роблять графові методи практично непридатними для великих зображень.

1.7 Сегментація зображень на основі кластерного аналізу даних та штучних нейронних мереж

Кластерний аналіз (англ. data clustering) - завдання розбиття заданої вибірки об'єктів (ситуацій) на підмножини, які називаються кластерами, так, щоб кожен кластер складався із схожих об'єктів, а об'єкти різних кластерів істотно відрізнялися. Завдання кластеризації відноситься до статистичної обробки, а також до широкого класу задач навчання без учителя. Кластерний аналіз - це багатовимірна статистична процедура, що виконує збір даних, що містять інформацію про вибірку об'єктів, і потім впорядковує об'єкти в порівняно однорідні групи (кластери) (Q-кластеризація, або Q-техніка, власне кластерний аналіз). Кластер - група елементів, характеризуються загальною властивістю, головна мета кластерного аналізу - знаходження груп схожих об'єктів у вибірці. Спектр застосувань кластерного аналізу дуже широкий: його використовують в антропології, археології, медицині, інформаційних технологіях, психології, хімії, біології, державному управлінні, філології, маркетингу, соціології та інших дисциплінах. Проте універсальність застосування призвела до появи великої кількості несумісних термінів, методів і підходів, що ускладнюють однозначне використання та

несуперечливу інтерпретацію кластерного аналізу.

Кластерний аналіз виконує такі основні завдання:

- розробка типології та класифікації;
- дослідження корисних концептуальних схем групування об'єктів;
- породження гіпотез на основі дослідження даних;
- перевірка гіпотез або дослідження для визначення того, чи типи (групи), виділені тим чи іншим способом, присутні у наявних даних.

Незалежно від предмета вивчення застосування кластерного аналізу передбачає такі етапи:

- відбір вибірки для кластеризації;
- визначення множини змінних, по яких будуть оцінюватися об'єкти у вибірці;
- обчислення значень тієї чи іншої міри подібності між об'єктами;
- застосування методу кластерного аналізу для створення груп схожих об'єктів;
- перевірка достовірності результатів кластерного рішення.

Кластерний аналіз пред'являє такі вимоги до даних: по-перше, показники не повинні корелювати між собою, по-друге, показники повинні бути безрозмірними, по-третє, їх розподіл має бути близько до нормального; по-четверте, показники повинні відповідати вимогу «стійкості», під якою розуміється відсутність впливу на їх значення випадкових факторів; по-п'яте, вибірка повинна бути однорідна, не містити «викидів». Якщо кластерному аналізу передують факторний аналіз, то вибірка не потребує «ремонту» - викладені вимоги виконуються автоматично самою процедурою факторного моделювання (є ще одна перевага - z-стандартизація без негативних наслідків для вибірки; якщо її проводити безпосередньо для кластерного аналізу, вона може спричинити за собою зменшення чіткості поділу груп). В іншому випадку вибірку треба коригувати.

Тепер виникає питання стійкості прийнятого кластерного рішення. По суті, перевірка стійкості кластеризації зводиться до перевірки її

достовірності. Тут існує емпіричне правило - стійка типологія зберігається при зміні методів кластеризації. Результати ієрархічного кластерного аналізу можна перевіряти ітеративним кластерним аналізом за методом k-середніх. Якщо порівнювані класифікації груп респондентів мають частку збігів більше 70% (більше 2 / 3 збігів), то кластерне рішення приймається.

Перевірити адекватність рішення, не вдаючись до допомоги іншого виду аналізу, не можна. Принаймні, в теоретичному плані ця проблема не вирішена. У класичній роботі Олдендерфера і Блешфілда «Кластерний аналіз» докладно розглядаються і в підсумку відкидаються додаткові п'ять методів перевірки стійкості:

- кофенетична кореляція - не рекомендується і обмежена у використанні;
- тести значущості (дисперсійний аналіз) - завжди дають значимий результат;
- методика повторних (випадкових) вибірок, що, тим не менш, не доводить обґрунтованість рішення;
- тести значущості для зовнішніх ознак придатні тільки для повторних вимірювань;
- методи Монте-Карло дуже складні і доступні тільки досвідченим математикам.

Типологія завдань кластеризації описується таким типом вхідних даних:

- ознаковий опис об'єктів. Кожен об'єкт описується набором своїх характеристик, які називаються ознаками. Ознаки можуть бути числовими або нечисловими;
- матриця відстаней між об'єктами. Кожен об'єкт описується відстанями до всіх інших об'єктів навчальної вибірки.

Цілі кластеризації:

- розуміння даних шляхом виявлення кластерної структури. Розбиття вибірки на групи схожих об'єктів дозволяє спростити подальшу обробку

даних і прийняття рішень, застосовуючи до кожного кластеру свій метод аналізу (стратегія «розділяй і володарюй»);

- стиснення даних. Якщо вихідна вибірка надлишково велика, то можна скоротити її, залишивши по одному найбільш типовому представникові від кожного кластера.

- виявлення новизни (англ. novelty detection). Виділяються нетипові об'єкти, які не вдається приєднати до жодного з кластерів.

У першому випадку число кластерів намагаються зробити поменше. У другому випадку важливіше забезпечити високу ступінь подібності об'єктів усередині кожного кластера, а кластерів може бути скільки завгодно. У третьому випадку найбільший інтерес представляють окремі об'єкти, які не вписуються ні в один із кластерів.

У всіх цих випадках може застосовуватися ієрархічна кластеризація, коли великі кластери дробляться на більш дрібні, ті в свою чергу дробляться ще дрібніше, і т. д. Такі завдання називаються завданнями таксономії. Результатом таксономії є деревоподібна ієрархічна структура. При цьому кожен об'єкт характеризується перерахуванням всіх кластерів, яких він належить, звичайно від великого до дрібного. Класичним прикладом таксономії на основі схожості є Біноміальна номенклатура живих істот, запропонована Карлом Ліннеєм у середині XVIII століття. Аналогічні систематизації будуються в багатьох областях знання, щоб упорядкувати інформацію про велику кількість об'єктів.

Основна проблема методів кластеризації полягає в тому, що просторове розташування точок або не враховується зовсім, або враховується опосередковано (наприклад, використовуючи координати точки як одна з ознак). Тому зазвичай після кластеризації точок зображення проводять процедуру виділення зв'язних компонент. Методи кластеризації погано працюють на зображеннях із завадами: часто втрачають окремі точки регіонів, утворюється багато дрібних регіонів, і. т. п.

1.8 Сучасний стан і перспективи розвитку методів сегментації

Сегментація зображень посідає значне місце у системах розпізнавання та обробки інформації. Постійний розвиток технологій обробки зображень призвели до виникнення нових підходів до вирішення завдань сегментації зображень і застосування їх при вирішенні багатьох практичних завдань. На теперішній час існує значна кількість методів для проведення сегментації зображень. В залежності від типу зображення та його використання підходи до сегментації будуть різними. Вибір методу сегментації буде залежати від вимог до швидкості виконання сегментації та необхідності втручання користувача в процес сегментації. Але як свідчить аналіз багатьох експериментів, найкращі результати показують методи сегментації границь по контрастності на основі використання адаптивних масок. Цей підхід є найбільш перспективним у теперішній час.

1.9 Локалізація зображень

Сегментовані зображення самі по собі можуть розглядатися як результат локалізації. Проте у деяких системах є потреба обробляти зображення об'єкту з урахуванням його контексту (чи околу). У такій ситуації необхідно побудувати окіл сегментованої множини пікселів. Таке завдання можна вирішити декількома. Найпростіше – побудувати ϵ -окіл границі сегментованого об'єкту. Але на практиці як найчастіше є умова побудувати прямокутну область навколо об'єкту.

Для вирішення цієї задачі будемо використовувати метод головних компонент (МГК), який дозволяє знайти орієнтацію об'єкту у просторі з метою адекватної апроксимації області об'єкту прямокутником (рисунок 1.5). Детальніше розглянемо цей аспект у наступному розділі.

2 МЕТОДИ МАШИННОГО НАВЧАННЯ ДЛЯ СЕГМЕНТАЦІЇ ТА ЛОКАЛІЗАЦІЇ ЗОБРАЖЕНЬ

2.1 Класична модель сегментації

Значна частина завдань обробки інформації й аналізу даних пов'язана із зображеннями. Прикладами можуть служити обробка й аналіз даних дистанційного зондування Землі із супутників, об'ємні зображення об'єктів, отримані за допомогою голографічних пристроїв, результати застосування методів неруйнівних досліджень і контролю в промисловості, розробка “органів зору” роботів і сучасних медичних систем діагностики (рентгенографія, машинна томографія, ендоскопія й т.д.). Для успішного рішення завдань пошуку й ідентифікації об'єктів, визначення різного роду їхніх кількісних характеристик необхідно, щоб первинні зображення характеризувалися високою візуальною якістю, що втрачається через незадовільні умови одержання зображень, недосконалості систем передачі відеоінформації і її відображень, впливи перешкод і т.п.

Виділення контурів (виділення меж) – один із напрямів сегментації зображень, що широко використовується у теорії обробки зображень і комп'ютерного зору, частково з області пошуку об'єктів і виділення об'єктів, та ґрунтується на алгоритмах, які виділяють точки цифрового зображення, в яких різко змінюється яскравість або є інші види неоднорідності.

Основною метою виявлення різких змін яскравості зображення є фіксація контуру об'єктів зображення. Вони можуть відображати різні припущення про модель формування зображення, зміни в яскравості зображення можуть вказувати на: зміну глибини; зміну орієнтації поверхонь; зміну у властивостях матеріалу; відмінність у висвітленні сцени.

В ідеальному випадку результатом виділення контурів є набір зв'язаних кривих, що позначають межі об'єктів, граней і відбитків на поверхні, а також

криві, які відображають зміни положення поверхонь. Таким чином, застосування фільтра виділення меж до зображення може істотно зменшити кількість оброблюваних даних, через те що відфільтрована частина зображення вважається менш значущою, а найбільш важливі структурні властивості зображення зберігаються. Однак, на жаль, не завжди можливо виділити межі на зображенні в картинах реального світу середньої складності. Межі виділені з таких зображень часто мають такі недоліки, як фрагментованість (криві меж не з'єднані між собою), відсутність меж або наявність помилкових, що не відповідають досліджуваному об'єкту, меж.

Межі мають досить важливе значення в багатьох сферах застосування обробки зображень, особливо в системах машинного зору, які аналізують сцени штучних об'єктів при фіксованому освітленні.

Хоча деяка література розглядає виділення ідеальних східчастих меж, кордони на натуральному зображенні зазвичай не такі. На них зазвичай впливає один або кілька таких ефектів: фокусне розмиття через кінцеву глибини різкості зйомки, розмита півтінь від неточкових джерел світла, затінення гладких об'єктів, і тому багато дослідників використовують ступінчастий край, згладжений функцією Гауса (функція помилки), як найпростіше наближення моделі ідеального краю для моделювання розмитих меж у прикладних задачах.

Існує безліч підходів до виділення контурів (меж), але практично всі можна розділити на дві категорії: методи, засновані на пошуку максимумів, і методи, засновані на пошуку нулів. Методи, засновані на пошуку максимумів, виділяють межі за допомогою обчислення «сили краю», зазвичай вирази першої похідної, такого як величина градієнта, і потім пошуку локальних максимумів сили краю, використовуючи передбачуваний напрям межі, зазвичай перпендикуляр до вектора градієнта. Методи, засновані на пошуку нулів, шукають перетин осі абсцис шляхом обчислення другої похідної, зазвичай нулі лапласіана або нулі нелінійного диференціального виразу. В якості кроку передобробки до виділення меж

практично завжди застосовується згладжування зображення, зазвичай фільтром Гауса.

Канні вивчив математичну проблему отримання фільтра, оптимального за критеріями виділення, локалізації та мінімізації кількох відгуків одного краю. Він показав, що шуканий фільтр є сумою чотирьох експонент. Він також показав, що цей фільтр може бути добре наближений першою похідною Гаусіани. Канні ввів поняття Non-Maximum Suppression (придушення не-максимумів), яке означає, що пікселями меж оголошуються пікселі, в яких досягається локальний максимум градієнта в напрямку вектора градієнта. Хоча його робота була проведена на зорі комп'ютерного зору, детектор меж Канні до сих пір є одним з кращих детекторів. Крім особливих локальних випадків важко знайти детектор, який би працював істотно краще, ніж детектор Канні. Послідовність роботи детектора Канні складається з таких етапів.

Згладжування. Розмиття 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 порогами. Трасування області неоднозначності. Підсумкові межі визначаються шляхом придушення всіх країв, незв'язаних з певними (сильними) межами.

Детектор Канні-Деріше був виведений зі схожого математичного критерію, як і детектор Канні, хоча, відштовхуючись від іншої точки зору, він призвів до набору рекурсивних фільтрів для згладжування зображення замість експоненційних фільтрів і фільтрів Гауса.

Після того, як ми знайшли силу межі (зазвичай - величину градієнта), наступний крок - застосувати поріг, щоб вирішити, знаходиться чи ні межа в даній точці зображення. Чим менший поріг, тим більше пікселів буде перебувати на межі, але тим більш сприйнятливим до завад стане результат, виділяючи зайві дані зображення. Навпаки, високий поріг може пропустити

слабкі краї або отримати межу фрагментами.

Якщо поріг застосовується просто до зображення величини градієнта, отримані межі будуть товстими і буде потрібно деяка постобробка, що робить край тонким і точним. Якщо ж виділити межі за допомогою Non-Maximum Suppression, кордони будуть тонкими за визначенням і їх можна буде поєднати в полігони процедурою з'єднання країв (простежування межі). На дискретній сітці етап придушення немаксимумів може бути реалізований за допомогою оцінки напрямку градієнта, використовуючи перші похідні, округлення напрямлення на значення з кроком 45 градусів і, нарешті, порівнянні значень градієнта в отриманому напрямку градієнту.

Традиційним підходом до вирішення проблеми знаходження відповідного порогу є пороги «з запізненням». Метод використовує декілька порогів. Ми використовуємо верхній поріг, щоб знайти точку початку межі. Після того, як ми отримали стартову точку, ми відстежуємо кордон, точка за точкою, поки значення сили краю вище нижнього порогу. Цей алгоритм має на увазі, що межі - це швидше за все безперервні криві, і дозволяє нам простежувати слабкі ділянки меж без допущення того, що всі значні точки на зображенні будуть позначені як край. Однак, у нас все ще є проблема вибору відповідних значень порогів для цього методу, так як оптимальні параметри можуть змінюватися від зображення до зображення.

Уточнення контуру (межі) - процес, який робить контури тонкими, видаляючи небажані хибні точки, які з'являються на межі. Ця техніка застосовується вже після того, як зображення було згладжено (використовуючи фільтр Гауса), був застосований оператор меж (як один з описаних вище) для обчислення сили краю і після того, як межі були очищені, використовуючи відповідні пороги. Цей метод видаляє всі небажані точки і при акуратному застосуванні видає контури товщиною в один піксель.

Плюси: різкі й тонкі межі дозволяють підвищити ефективність розпізнавання об'єктів; при використанні трансформації Хафа для виявлення

прямих або еліпсів, тонкі межі дають значно кращі результати; якщо межа являє собою межа деякої області, тонкі межі дозволяють обчислити такі параметри, як периметр, без якоїсь складної арифметики.

Вказані підходи зазвичай використовують для реалізації завдання по виділення границь об'єктів комп'ютерних зображень, як правило, наступним чином: а) переведення зображення із моделі кольорів RGB в YUV; б) видалення можливих завад зображення шляхом фільтрації; в) згладження зображення з використанням фільтрів Гауса; г) просторове диференціювання зображення; г) пошук локальних максимумів та мінімумів; д) зв'язування контурів (пошук границь об'єктів від знайдених максимумів та мінімумів); е) кластеризація зображення; є) накладання отриманих границь об'єктів на початкове зображення; і) виведення сегментованого зображення.

Відмінність такої моделі у тому, що введено процедуру фільтрації зображення від можливих завад та введено процедуру кластеризації об'єктів зображення з метою спрощення подальшого використання сегментованого зображення у системах розпізнавання, а також вилучені процедури із подавленням НЕ-максимумів.

2.2 Виділення границь за допомогою операторів

Одним із підходів до пошуку однорідних областей шляхом виділення контурів є використання операторів. Таким є оператори Собеля, Преввита (рисунок 1.1) які використовуються в області обробки зображень для виділення контурів об'єктів на зображенні. Ці оператори досить часто застосовують в алгоритмах виділення меж. По суті, це дискретний диференціальний оператор, який обчислює наближене значення градієнта яскравості зображення. Результатом застосування оператора Собеля в кожній точці зображення є або вектор градієнта яскравості в цій точці, або його норма. Оператор Собеля заснований на згортку зображення невеликими сепарабельним цілочисельними фільтрами у вертикальному та

горизонтальному напрямках, тому його відносно легко обчислювати. З іншого боку, використовувана ним апроксимація градієнта досить груба, особливо це позначається на високочастотних коливань зображення.

Якщо простіше, то оператор обчислює градієнт яскравості зображення в кожній точці. Так знаходиться напрям найбільшого збільшення яскравості і величина її зміни в цьому напрямі. Результат показує, наскільки «різко» або «плавно» змінюється яскравість зображення в кожній точці, а отже, ймовірність знаходження точки на межі, а також орієнтацію межі. На практиці, обчислення величини зміни яскравості (імовірності належності до грані) надійніше і простіше в інтерпретації, ніж розрахунок напрямку.

Математично, градієнт функції двох змінних для кожної точки зображення (якої і є функція яскравості) - двовимірний вектор, компонентами якого є похідні яскравості зображення по горизонталі й вертикалі. У кожній точці зображення градієнтний вектор орієнтований у напрямку найбільшого збільшення яскравості, а його довжина відповідає величині зміни яскравості. Це означає, що результатом оператора Собеля в точці області постійної яскравості буде нульовий вектор, а в точці, що лежить на кордоні областей різної яскравості - вектор, що перетинає межу у напрямку збільшення яскравості.

Оскільки функція яскравості відома тільки в дискретних точках, ми не можемо визначити похідні до тих пір, поки ми покладемо яскравість безперервною функцією, яка проходить через ці точки. З цієї додаткової передумовою похідну безперервної функції яскравості можна обчислити як від функції, з якої взято виміри - точки зображення. Виявляється, що похідні в будь-якій окремій точці є функції яскравості від всіх точок зображення. Однак наближення їх похідних можна визначити з більшою чи меншою мірою точності.

Оператор Собеля представляє собою більш наочне наближення градієнта зображення, але він досить якісний для практичного застосування у багатьох завданнях. Точніше, оператор використовує значення інтенсивності

тільки в околиці 3×3 кожного пікселя для отримання наближення відповідного градієнта зображення, і використовує тільки цілочисельні значення вагових коефіцієнтів яскравості для оцінки градієнту.

Як випливає з визначення, оператор Собеля можна реалізувати простими технічними та програмними засобами: для наближення вектор-градієнта потрібні тільки вісім пікселів навколо точки зображення і цілочисельна арифметика. Більше того, обидва дискретних фільтра, описаних вище, можна розділити на дві групи. Роздільність цих обчислень може призвести до зменшення арифметичних дій з кожним пікселів.

Результат застосування оператора Собеля є двовимірна карта градієнта для кожної точки. Її можна обробити і показати як картинку, на якій ділянки з великою величиною градієнта (в основному, грані) будуть видні як білі лінії. Оператор Собеля згладжує паразитні ефекти на зображенні, що викликаються чисто центрально-диференціальним оператором, але не має повної обертальної симетрії.

Оператор Превітт – метод виділення меж в обробці зображень, який обчислює максимальний відгук на множині ядер згортки для знаходження локальної орієнтації межі в кожному пікселі. Він був створений доктором Джудіт Превітт (Judith Prewitt) для виявлення меж медичних зображень.

Для цієї операції використовуються різні ядра. З одного ядра можна отримати вісім, переставляючи коефіцієнти по колу. Кожен результат буде чутливий до напрямку межі від 0 до 315 з кроком в 45, де 0 відповідає вертикальній межі. Максимальна відповідь кожного пікселя є значення відповідного пікселя у вихідному зображенні. Значення його лежать між 1 і 8, в залежності від номера ядра, що дав найбільший результат.

Цей метод виділення меж також називається підстановкою шаблонів меж (англ. edge template matching), оскільки зображенню зіставляється з набором шаблонів, і кожен представляє деяку орієнтацію межі. Величина і орієнтація межі в пікселі тоді визначається шаблоном, який краще за всіх відповідає локальній околиці пікселя.

Детектор меж Превітт є підходящим способом для оцінки величини та орієнтації межі. У той час як детектор з диференціальним градієнтом потребує трудомісткого обчислення оцінки орієнтації за величинами у вертикальному та горизонтальному напрямках, детектор меж Превітт дає напрямок прямо з ядра з максимальним результатом. Набір ядер обмежений 8 можливими напрямками, проте досвід показує, що більшість прямих оцінок орієнтації теж не дуже точні. З іншого боку, набір ядер потребує 8 згорток для кожного пікселя, тоді як набір ядер градієнтного методу вимагає тільки 2: чутливих по вертикалі і по горизонталі. Результат для зображення потужності меж дуже схожий в обох методів, якщо в них використовуються ті ж ядра згортки.

Оператор Собеля використовує два ядра 3×3 , згортаючи початкове зображення для обчислення наближених значень похідних: одне по горизонталі і одне по вертикалі.

Застосування оператора Превітт дає кращі результати тоді, коли на зображенні необхідно виділити межу з переважаючим напрямком орієнтації. Розглянуті оператори будуть використані при створенні програми для виділення контурів комп'ютерних зображень.

З метою поліпшення якості сегментації за рахунок підвищення чутливості до не контрастних перепадів яскравості будемо використовувати лапласіан гауссіана. Лапласіан зображення в пікселі має вигляд

$$L(x, y) = \frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2}; \quad (2.1)$$

Для найбільш поширених згорток 3×3 матриця лапласіану представлена на рисунку 2.1.

На основі використання цього оператору [1] вводиться до розгляду оператор лапласіан гауссіана

$$\text{LoG}(x, y) = \frac{-1}{\pi\sigma^4} \left[1 - \frac{x^2 + y^2}{2\sigma^2} \right] + e^{-\frac{x^2 + y^2}{2\sigma^2}}; \quad (2.2)$$

геометрична інтерпретація якого наведено на рисунку 2.2.

0	-1	0
-1	4	-1
0	-1	0

-1	-1	-1
-1	8	-1
-1	-1	-1

Рисунок 2.1 – Найбільш поширені дискретні апроксимації оператора Лапласа

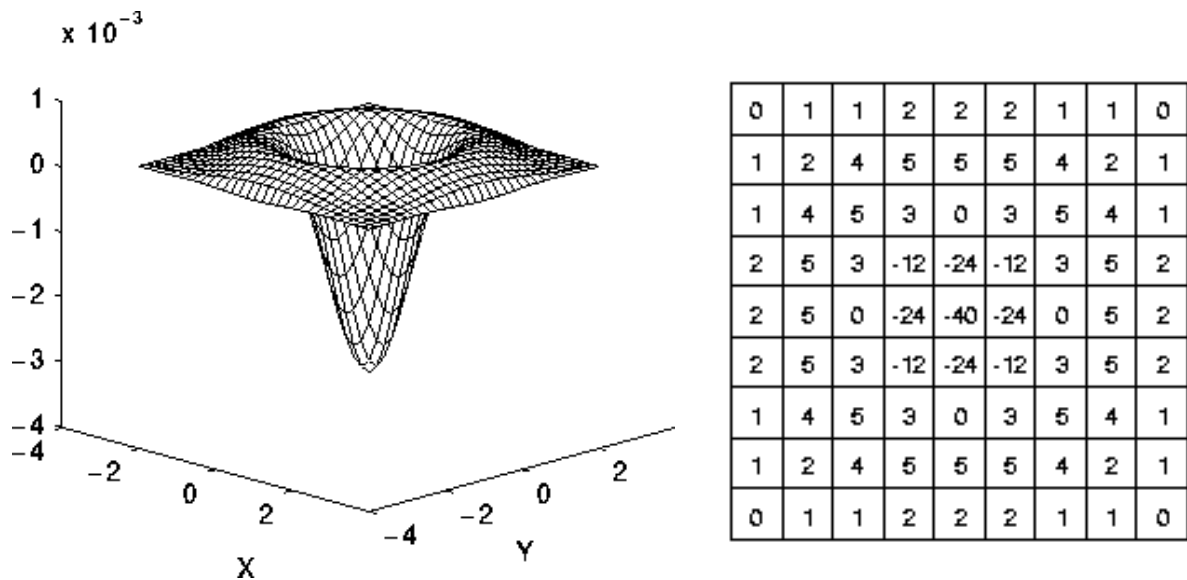


Рисунок 2.2 – Геометрична інтерпретація оператора LOG і його матриця перетворень, де $\sigma = 1.4$

Обраний оператор дозволяє дуже якісно виділяти граничні профілі зображень, у тому слабо контрастні профілі, завдяки підвищенню різкості границь таким чином, як це наведено на рисунку 2.3. Приклади застосування оператора наведено на рисунку 2.4.

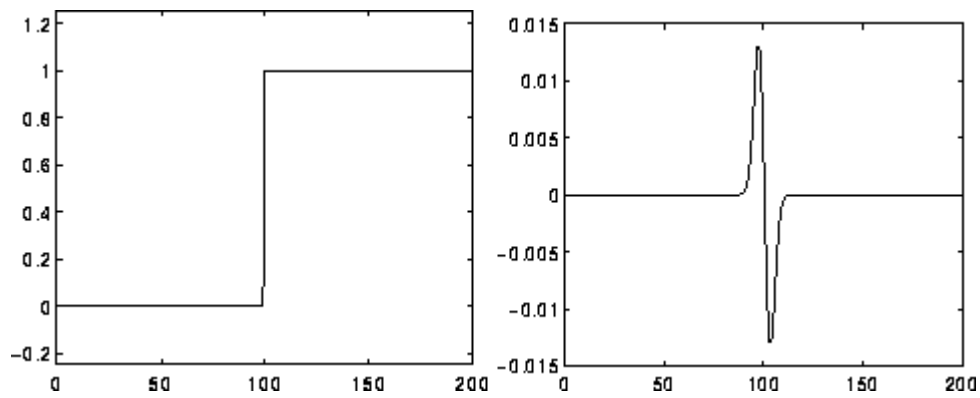


Рисунок 2.3 – Формування граничного переходу оператором LoG

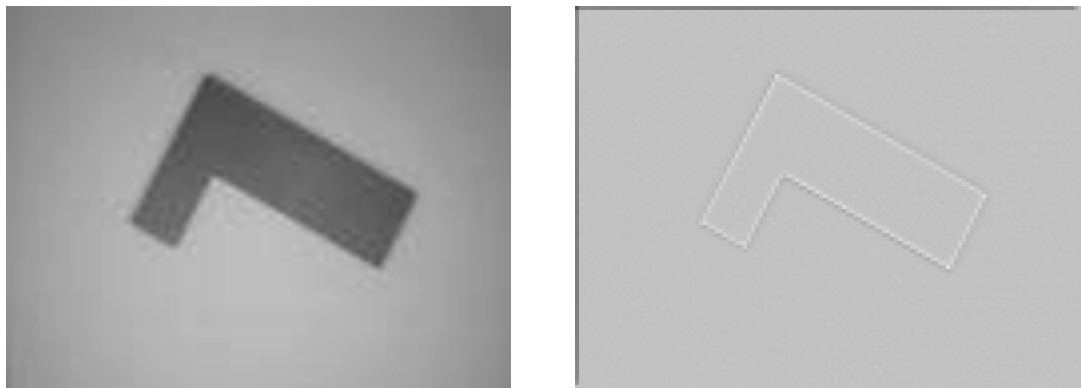


Рисунок 2.4 – Застосування оператора LoG

2.3 Розробка алгоритму сегментації зображення

Знаходження границі зображення (рисунок 2.5) починається з знаходження граничних пікселів. Але для високої ефективності контурної сегментації (для мінімізації рівнів часткової і хибної сегментації) до її початка експерт будує границі зображень (рисунок 2.6) з метою настроювання параметрів моделі (порогів) таким чином, щоб мінімізувати відхилення сегментованої границі і талону, який побудував експерт.

Таким чином, маючи еталон (рисунок 2.6) ми спроможні ітераційно варіювати порогові і оптимально настроюватися на еталон (рисунок 2.7). Далі для сегментації використовують ці оптимальні значення порогів, які дають мінімальні рівні часткової і хибної сегментації (рисунок 2.8).

Після отримання первинних результатів сегментації з використанням оператора лапласіан гауссіана другий крок – це застосування методу водорозділів для отримання замкнутих контурів, які відповідають об'єктам на фотографії. Цей метод доцільно застосовувати тому, що об'єкти розташовані на однорідному фоні.

Після цього застосовуємо метод хвиль для отримання зв'язних областей, які відповідають об'єктам.

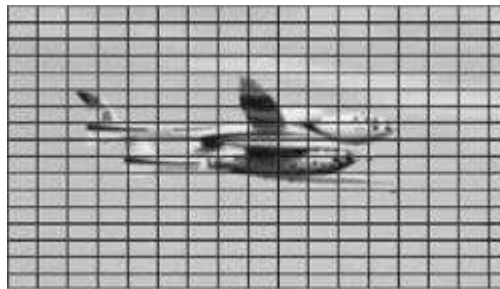


Рисунок 2.5 – Вихідне зображення



Рисунок 2.6 – Експертне оцінювання границь зображень



Рисунок 2.7 – Ілюстрація процесу настроювання порогів



Рисунок 2.8 – Сегментація границі

Наші зображення завжди мають завади і шуми. Наявність яких приводить до хибної сегментації фрагментів вихідного зображення (рисунок 2.9). Тому, після застосування методу хвиль застосовуємо геометричну фільтрацію з метою усунення усіх зв'язних компонент, які не є об'єктами. Для такої фільтрації достатньо використовувати параметр площини зображення (рисунок 2.10).



Рисунок 2.9 – Сегментація границі з наявністю об'єктів і хибних фрагментів вихідного зображення

Наступний крок – застосування морфологічної фільтрації (операція дилатації – зростання світлих областей; ядро 3×3) для видалення «дірок» з внутрішці об'єктів (рисунок 2.9 – 2.11).

Аналіз результатів обчислювального експерименту дозволяє зробити висновок про те, що запропонована модель (алгоритм) адаптивної сегментації є обчислювально ефективною і здатною адекватно сегментувати контур об'єкта нерегулярного вигляду в умовах існування завад, а також в умовах істотних варіацій яскравості об'єктів і фону.

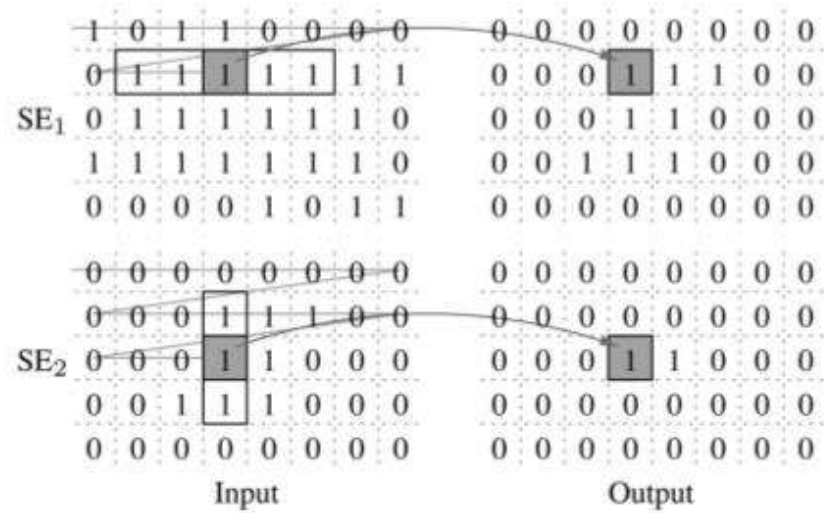


Рисунок 2.10 – Морфологічна фільтрація з усуненням дірок (фільтр Open)

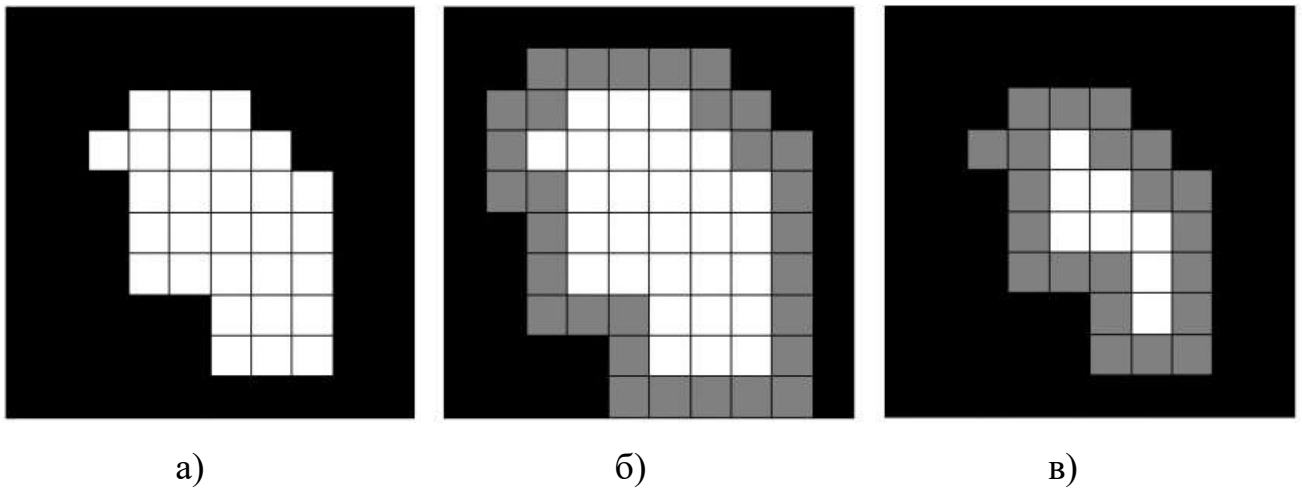


Рисунок 2.11 – Вихідне зображення (а), дилатація (б) і ерозія (в) зображення

2.4 Локалізація з використанням методу головних компонент

Оскільки двовимірний простір ознак являє типовий випадок, розглянемо спочатку цей метод для простору \mathbb{R}^2 з координатами (r_1, r_2) з метою отримання формул, використання яких дозволяє підвищити його обчислювальну ефективність.

Для отримання більш компактного покриття значень ознак $X = \{x_i = (x_{1i}, x_{2i})\}_{i=1, \dots, n}$, заданих у вихідній системі координат (r_1, r_2) , яка визначається базисними векторами $Y = (y_1, y_2)$, шукається новий базис $Z = (z_1, z_2)$, в якому аппроксимируючий паралелепіпед матиме меншу площу. При $k > 2$ ортогональні вектори, що визначають новий базис, знаходяться послідовно, за критерієм максимального розсіювання вздовж осей; в двовимірному ж випадку нова система координат в базисі Z з точністю до орієнтації осей визначається першою головній компонентою $z_1 = (z_{11}, z_{12})$, яка максимізує вибірккову дисперсію проекції даних (рис. 2.15), а друга компонента $z_2 = (z_{21}, z_{22})$, з точністю до орієнтації, найбільш просто визначається з умови ортогональності, наприклад, $z_{21} = -z_{12}$, $z_{22} = z_{11}$.

Перехід до базису $Z = (z_1, z_2)$ підвищує адекватність геометричній моделі безлічі X за рахунок виключення з розгляду областей, значення ознак в яких не відповідають стандартам.

Застосування методу головних компонент передбачає попереднє центрування даних. Для цього знаходимо незміщене положення початку системи координат $O' = (\bar{x}_1, \bar{x}_2)$

$$\bar{x}_1 = \frac{1}{n} \sum_{i=1}^n x_{1i}, \quad \bar{x}_2 = \frac{1}{n} \sum_{i=1}^n x_{2i}, \quad (2.3)$$

і центруємо дані щодо початку координат $O' = (\bar{x}_1, \bar{x}_2)$

$$x'_i = (x_{1i} - \bar{x}_1, x_{2i} - \bar{x}_2), \quad (i = 1, 2, \dots, n). \quad (2.4)$$

Для спрощення індексації покладемо, що координати точок, складяючих безліч X , задані саме в цій системі координат.

Знаходження компонент перетворення до головних осей (z_1, z_2) зводиться до знаходження власних значень $\lambda = (\lambda_1, \lambda_2)$ і власних векторів $V = (v_1, v_2)$ ковариаційної матриці ознак

$$C = \begin{bmatrix} a & c \\ c & b \end{bmatrix}, \quad (2.5)$$

де діагональні елементи визначають дисперсії ознак

$$a = \sigma_{11} = \frac{1}{n} \sum_{i=1}^n (x_{1i} - \bar{x}_1)^2, \quad b = \sigma_{22} = \frac{1}{n} \sum_{i=1}^n (x_{2i} - \bar{x}_2)^2, \quad (2.6)$$

а інші елементи утворюють їх взаємні ковариації

$$c = \sigma_{12} = \sigma_{21} = \frac{1}{n} \sum_{i=1}^n (x_{1i} - \bar{x}_1) \cdot (x_{2i} - \bar{x}_2). \quad (2.7)$$

Оскільки для зосереджених даних значення середніх \bar{x}_1, \bar{x}_2 дорівнюють нулю, останні вираження приймуть вид

$$\sigma_{11} = \frac{1}{n} \sum_{i=1}^n (x_{1i})^2, \quad \sigma_{22} = \frac{1}{n} \sum_{i=1}^n (x_{2i})^2, \quad (2.8)$$

$$\sigma_{12} = \sigma_{21} = \frac{1}{n} \sum_{i=1}^n x_{1i} \cdot x_{2i}. \quad (2.9)$$

Якщо E - одинична матриця, то власні значення $\lambda = (\lambda_1, \lambda_2)$ та вектори $V = (v_1, v_2)$ матриці c знаходяться з рівняння

$$(C - \lambda \cdot E) \cdot v = 0, \quad (2.10)$$

яке представляє собою однорідну систему лінійних рівнянь, нетривіальні рішення якої є шуканими власними векторами матриці.

В явному вигляді рівняння (2.10) представляється в матричному вигляді так

$$\left(\begin{bmatrix} a & c \\ c & b \end{bmatrix} - \lambda \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) \cdot \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} a - \lambda & c \\ c & b - \lambda \end{bmatrix} \cdot \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = 0, \quad (2.11)$$

де v_1, v_2 позначають компоненти власного вектора v_ξ , $\xi \in \{1,2\}$, або системою лінійних рівнянь

$$\begin{cases} (a - \lambda) \cdot v_1 + c \cdot v_2 = 0, \\ c \cdot v_1 + (b - \lambda) \cdot v_2 = 0, \end{cases} \quad (2.12)$$

яке визначає характеристичне рівняння для власних значень виду

$$\lambda^2 - (a + b) \cdot \lambda + (a \cdot b - c^2) = 0,$$

з дискримінантом

$$D = (a - b)^2 + 4c^2 > 0,$$

коріння якого мають вигляд

$$\begin{cases} \lambda_1 = \frac{(a + b) + \sqrt{D}}{2}, \\ \lambda_2 = \frac{(a + b) - \sqrt{D}}{2}. \end{cases} \quad (2.13)$$

Використання цих значень в системі (2.12), визначає два підпростору власних векторів, що відповідають значенням λ_1 і λ_2 , в кожному з яких вектори лінійно залежні, а значить для побудови нового базису $Z = (z_1, z_2)$ може вибиратися будь-який з них. Для отримання ортонормованій системи виберемо одиничні вектори. При цьому перший власний вектор v_1 матриці C відповідає максимальному власному значенню λ_1 , а другий – значенню λ_2 , або ортогональним йому одиничним векторам.

Таким чином, підставляючи значення λ_1 з (2.12) і поповнюючи систему умовою нормування

$$\sqrt{v_{11}^2 + v_{12}^2} = 1,$$

отримаємо, що в старих координатах нормований вектор головної компоненти, що визначає вісь z_1 , виражається в такий спосіб

$$\begin{cases} z_{11} = \sqrt{\frac{c^2}{c^2 + (\lambda_1 - a)^2}}, \\ z_{12} = \sqrt{\frac{(\lambda_1 - a)^2}{c^2 + (\lambda_1 - a)^2}}. \end{cases}$$

Для отримання координат вектора, що визначає вісь z_2 , можна використовувати друге власне значення λ_2 , або, що вимагає меншого обсягу обчислень, можна повернути вектор z_1 на кут $\pi/2$; в цьому випадку отримуємо

$$\begin{cases} z_{21} = -z_{12}, \\ z_{22} = z_{11}. \end{cases}$$

Якщо при цьому одна з компонент дорівнює нулю, то вихідний базис є оптимальним і ніяких дій робити не потрібно, оскільки вихідні осі і представляють головні компоненти. Тому далі вважаємо, що $z_{11} \neq 0$ і $z_{12} \neq 0$. В цьому випадку, оскільки обидва базису $Y = (y_1, y_2)$ і $Z = (z_1, z_2)$ ортонормовані, їх компоненти зв'язані співвідношеннями

$$\begin{cases} z_1 = \cos \alpha y_1 + \sin \alpha y_2, \\ z_2 = -\sin \alpha y_1 + \cos \alpha y_2; \\ \\ y_1 = \cos \alpha z_1 - \sin \alpha z_2, \\ y_2 = \sin \alpha z_1 + \cos \alpha z_2, \end{cases}$$

де

$$\begin{cases} \cos \alpha = z_{11}, \\ \sin \alpha = z_{12}. \end{cases} \quad (2.14)$$

Відповідно, координати (x'_{1i}, x'_{2i}) і (x_{1i}, x_{2i}) точки $\mathbf{x}_i = (x_{1i}, x_{2i})$ в новому і старому базисах пов'язані такими співвідношеннями

$$\begin{cases} x'_{1i} = \cos \alpha \cdot x_{1i} + \sin \alpha \cdot x_{2i}, \\ x'_{2i} = -\sin \alpha \cdot x_{1i} + \cos \alpha \cdot x_{2i}, \end{cases}$$

а кут повороту, якщо потрібно, можна знайти з (2.14).

3 РЕАЛІЗАЦІЯ ТА ТЕСТУВАННЯ МОДЕЛІ

3.1 Загальна модель сегментації і локалізації об'єктів

Відповідно до запропонованої моделі перш за все будемо використовувати модель контурної сегментації для знаходження границь об'єктів на цифровій фотографії (рисунок 3.1).

Після цього будемо застосовувати до контурів хвильовий метод задля їх зв'язування у окремі зв'язні компоненти границі. При цьому будемо виконувати трасування границі для того, щоб отримувати лише зовнішню границю об'єктів без гілок і петель. Щоб визначити однозв'язну область об'єкту, який розглядається.

Зазначимо, що у багатьох задачах цього достатньо і для локалізації об'єктів на знімку.

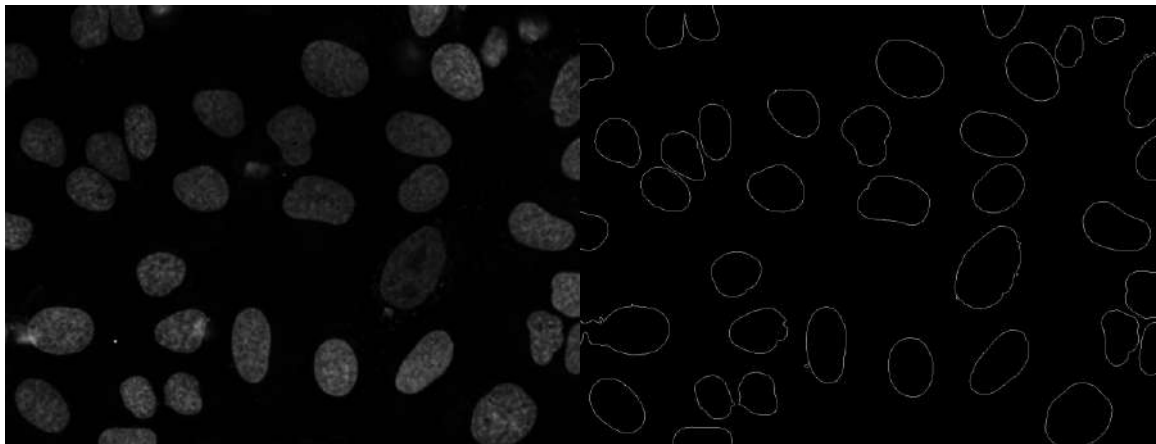


Рисунок 3.1 – Вихідна фотографія і контури об'єктів на ній

Якщо це не так, тоді до кожної такої зв'язної компоненти будемо застосовувати метод головних компонент, яких у задачах комп'ютерного зору використовують задля компенсації пороту об'єкту (рисунок 3.2), наприклад, в системах розпізнавання обличчя людини.

Застосування цього методу дозволяє адекватно визначити орієнтацію об'єкту у просторі. Після цього ми можемо визначити максимальні і мінімальні координати пік селів об'єкту по осям, знайти відповідні центри і визначити адекватні параметри прямокутника (рисунок 3.2).

Після цього координати такого прямокутника будуть однозначно визначати область локалізації об'єкту на знімку.

Після цього задачу локалізації у тому, чи іншому сенсі можна вважати вирішеною.



Рисунок 3.2 – Розпізнавання обличчя в системі

```

1 import matplotlib.pyplot as plt
2
3 from sklearn.model_selection import train_test_split
4 from sklearn.datasets import fetch_lfw_people
5 from sklearn.metrics import classification_report
6 from sklearn.decomposition import PCA
7 from sklearn.neural_network import MLPClassifier
8
9
10 # Load data
11 lfw_dataset = fetch_lfw_people(min_faces_per_person=100)
12
13 _, h, w = lfw_dataset.images.shape
14 X = lfw_dataset.data
15 y = lfw_dataset.target
16 target_names = lfw_dataset.target_names
17
18 # split into a training and testing set
19 X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3)
  
```

Рисунок 3.3 – Лістинг: Обробка даних е системи Face Recognition

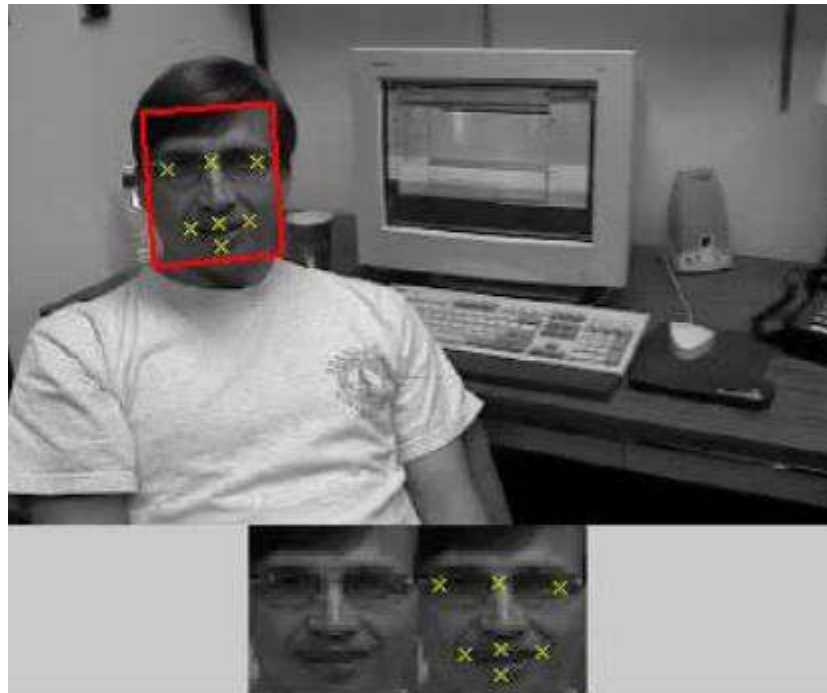


Рисунок 3.4 – Поворот зображення в системі розпізнавання з метою компенсації повороту обличчя

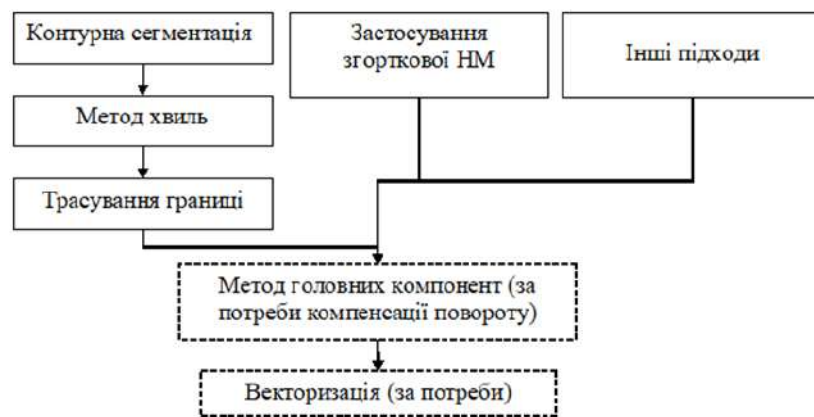


Рисунок 3.5 – Загальна модель локалізації об'єкту

Оскільки відтепер ми знаємо кут повороту зображення, після локалізації при виникненні потреби ми завжди можемо виконати нормалізацію і компенсувати поворот на цей кут (рисунок 3.4). Загальна модель локалізації об'єкту на зображенні наведено на рисунку 3.5.

ВИСНОВКИ

Проведено аналіз існуючих методів класифікації та кластеризації даних з використанням машинного навчання. аналіз існуючих методів класифікації та кластеризації даних. Проведено аналіз існуючих штучних нейронних мереж.

Відповідно завданню проведено аналіз сучасного стану питання, досліджені існуючі моделі і підходи до сегментації і локалізації об'єкту на цифровому зображенні з використанням методів машинного навчання. Побудовано відповідну загальну модель локалізації, з використанням якої проведений експеримент.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Гаптельманов А. В., Міхаль О. П., Щепка О. О. Підвищення енергозбереження бездротових сенсорних мереж з використанням методів машинного навчання // Системи управління навігації та зв'язку. Випуск 2 (72). Національний університет «Полтавська політехніка імені Юрія Кондратюка». Полтава 2023 р. С. 89-95.
2. Форсайт Д. А., Понс Дж. Комп'ютерний зір. Сучасний підхід - М.: Вільямс, 2004. - 928 с. - ISBN 5-8459-0542-7.
3. Шапіро Л., Штокман Дж. Комп'ютерний зір. – М.: Біном, 2009. – 763с.
4. Арлазаров В.Л. Сегментація об'єктів малого розміру на кольорових зображення / В.Л. Арлазаров, М.Д. Казанов // Програмування: Відділення математичних наук РАН. – 2008. – №3. – С. 65-76. - ISSN 0132-3474.
5. Горелік А.Л., Скрипкін В.А. Методи розпізнавання. - М: ВШ, 1989.- 256с.
6. Р. Гонсалес, Р. Вудс. Цифрова обробка зображень. - Москва: Техносфера, 2005. - 1072с.
7. Методи комп'ютерної обробки зображень. / За ред. В. А. Сойфер. - М. Фізматліт, 2003. - 784 с. - ISBN 5-9221-0270-2.
8. Потапов А. А., Пахомов А. А., Нікітін С. А., Гуляєв Ю. В. Нові методи обробки зображень. - М.: Фізматліт, 2008. - 496 с.
9. Рудаков П.І., Сафонов В.І. Обробка сигналів та зображень. - М: ДІАЛОГ-МІФІ, 2000. – 416 с.
10. Сойфер В.А. Комп'ютерна обробка зображень. Частина 1. Математичні моделі // Соросівський освітній журнал, 1996, №2 – с.5-28
11. Lucchese and S.K. Mitra “Color Image Segmentation: A State-of-the-Art Survey”, 2001. - pp. 1007-1019.
12. David Martin, Charless Fowlkes, Doron Tal, Jitendra Malik. A Database

of Human Segmented Natural Images and its Application to Evaluating Segmentation Algorithms and Measuring Ecological Statistics. Department of Electrical Engineering and Computer Sciences. - Pattern Recognition, Vol. 23, No 8, 2003.