

**НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені ІГОРЯ СІКОРСЬКОГО»**

**Радіотехнічний факультет
Радіотехнічних пристроїв та систем**

До захисту допущено:

Завідувач кафедри

_____ Сергій ЖУК

«__» _____ 20__ р.

Дипломна робота

на здобуття ступеня бакалавра

за освітньо-професійною програмою

«Радіотехнічні інформаційні технології»

спеціальності 172 «Телекомунікації та радіотехніка»

на тему: «Алгоритм виявлення рухомих об'єктів на зображенні»

Виконав:

студент III курсу, групи РТ-п-81
Моленко Йосип Рафатович

И. Ж

Керівник:

Старший викладач, кандидат технічних наук,
Вишневий Сергій Валерійович

Рецензент:

Доцент кафедри ТОР, кандидат технічних наук,
Гусева Олена Володимирівна

Засвідчую, що у цій дипломній роботі
немає запозичень з праць інших авторів
без відповідних посилань.

Студент И. Ж

Київ – 2021 року

Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»
Радіотехнічний факультет
Радіотехнічних пристроїв та систем

Рівень вищої освіти – перший (бакалаврський)

Спеціальність – 172 «Телекомунікації та радіотехніка»

Освітньо-професійна програма «Радіотехнічні інформаційні технології»

ЗАТВЕРДЖУЮ

Завідувач кафедри

_____ Сергій ЖУК

«__» _____ 2021 р.

ЗАВДАННЯ

на дипломну роботу студенту


Моленку Йосипу Рафатовичу

1. Тема роботи «Алгоритм виявлення рухомих об'єктів на зображенні», керівник роботи Вишневий Сергій Валерійович, старший викладач, затверджені наказом по університету від «18» травня 2021 р. №1205-с
2. Термін подання студентом роботи 07.06.2021 р.
3. Вихідні дані до роботи: Визначити актуальність та проаналізувати існуючі алгоритми виявлення рухомих об'єктів на зображенні.
4. Зміст пояснювальної записки:
Розділ 1: «АКТУАЛЬНІСТЬ ЗАДАЧІ ВИЯВЛЕННЯ РУХОМИХ ОБ'ЄКТІВ»
Розділ 2: «МЕТОДИ ТА АЛГОРИТМИ РОЗПІЗНАВАННЯ РУХОМИХ ОБ'ЄКТІВ»
Розділ 3: «РОЗРОБКА ТА ДОСЛІДЖЕННЯ АЛГОРИТМІВ РОЗПІЗНАВАННЯ РУХОМИХ ОБ'ЄКТІВ»
5. Дата видачі завдання 12 квітня 2021р.

Календарний план

№ з/п	Назва етапів виконання дипломної роботи	Термін виконання етапів роботи	Примітка
1	АКТУАЛЬНІСТЬ ЗАДАЧІ ВИЯВЛЕННЯ РУХОМИХ ОБ'ЄКТІВ	12.04.21 – 25.04.21	
2	МЕТОДИ ТА АЛГОРИТМИ РОЗПІЗНАВАННЯ РУХОМИХ ОБ'ЄКТІВ	24.04.21 – 16.05.21	
3	РОЗРОБКА ТА ДОСЛІДЖЕННЯ АЛГОРИТМІВ РОЗПІЗНАВАННЯ РУХОМИХ ОБ'ЄКТІВ	16.05.21 – 07.06.21	

Студент



Йосип Моленко

Керівник

Сергій Вишневий

**НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені ІГОРЯ СІКОРСЬКОГО»**

**Радіотехнічний факультет
Радіотехнічних пристроїв та систем**

Пояснювальна записка

до дипломної роботи

на тему: «Алгоритм виявлення рухомих об'єктів на зображенні»

Київ – 2021 року

ЗМІСТ

Перелік скорочень.....	9
ВСТУП	10
1 АКТУАЛЬНІСТЬ ЗАДАЧІ ВИЯВЛЕННЯ РУХОМИХ ОБ’ЄКТІВ.....	13
1.1. Актуальність теми.....	13
1.2. Проблеми задачі виявлення рухомих об’єктів.....	13
1.3. Сфери використання та можливості систем розпізнавання	15
1.4. Передобробка і постобробка зображень при виявленні рухомих об’єктів.....	25
1.4.1. Попередня обробка (передобробка).....	25
1.4.2 Постобробка	26
1.4.3 Математична морфологія.....	27
Висновки за розділом 1	29
2 МЕТОДИ ТА АЛГОРИТМИ РОЗПІЗНАВАННЯ РУХОМИХ ОБ’ЄКТІВ.....	30
2.1. Огляд методів виявлення рухомих об’єктів	30
2.2. Метод віднімання фону	34
2.3. Метод оптичного потоку.....	42
2.3.1. Оцінка оптичного потоку	42
2.3.2. Методи визначення оптичного потоку	43
Висновки за розділом 2	49
3 РОЗРОБКА ТА ДОСЛІДЖЕННЯ АЛГОРИТМІВ РОЗПІЗНАВАННЯ РУХОМИХ ОБ’ЄКТІВ	50
3.1. Розробка алгоритмів виявлення рухомих об’єктів	51

3.1.1. Розробка алгоритму віднімання фону.....	51
3.1.2. Розробка алгоритму віднімання фону з використанням суміші Гауссових розподілів (GMM).....	52
3.1.3. Розробка алгоритму Лукаса-Канаде для обчислення оптичного потоку	55
3.2. Дослідження алгоритмів виявлення рухомих об'єктів.....	56
Висновки за розділом 3	62
ВИСНОВКИ.....	63
Перелік посилань	65

АНОТАЦІЯ

Випускна кваліфікаційна бакалаврська робота: 67с., 14 рис., 1 табл., 31 джерел.

Тема: Алгоритм виявлення рухомих об'єктів на зображенні.

Об'єкт дослідження: процес проектування програми розпізнавання рухів об'єктів для системи відеонагляду.

Мета роботи: аналіз особливостей розробки автоматизованої моделі розпізнавання рухів об'єкту для системи відеонагляду.

Предмет дослідження: специфіка застосування алгоритму виявлення рухомих об'єктів на зображенні в системах відеонагляду.

Результати дослідження:

Досліджено можливості використання системи розпізнавання образів під час проектування автоматизованої моделі для камер відеоспостереження.

Висновки

В результаті здійснено розробку модулю розпізнавання рухів об'єкту для систем відеонагляду.

Ключові слова: РОЗПІЗНАВАННЯ ОБРАЗІВ, БАЗА ДАНИХ, ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ, ПРОЕКТУВАННЯ, ШТУЧНИЙ ІНТЕЛЕКТ, ССТV.

SUMMARY

Final qualifying bachelor's thesis: 67 pp., 14 figs., 1 tablet., 31 sources.

Topic: Algorithm for detecting moving objects in the image.

Object of research: the process of designing a program for recognizing the movements of objects for a video surveillance system.

Purpose: analysis of the features of the development of an automated model of object recognition for the video surveillance system.

Subject of research: the specifics of the application of the algorithm for detecting moving objects in the image in video surveillance systems.

Results of the research:

Possibilities of using the image recognition system during the design of an automated model for video surveillance cameras are investigated.

Conclusion

As a result, the object motion recognition module for video surveillance systems was developed.

Keywords: IMAGE RECOGNITION, DATABASE, SOFTWARE, DESIGN, ARTIFICIAL INTELLIGENCE, CCTV.

ПЕРЕЛІК СКОРОЧЕНЬ

БЗ – база знань;

АСВ – автоматизована система відеоспостереження;

ТРО – теорія розпізнавання образів;

ІІТ – інтелектуальні інформаційні технології;

ІБД – інтегрована база даних;

ЕОМ – електронно-обчислювальна машина;

ЕОД – електронний обмін даними;

ССТV – система відеоспостереження;

FPS – кількість кадрів в секунду;

BPP – кількість біт на піксель, також відоме як глибина кольору.

RGB – адитивна колірна модель;

GMM – Гауссова суміш розподілів;

MDPOF – Motion Detail Preserving Optical Flow Estimation.

ВСТУП

Актуальність дослідження. На сьогоднішньому етапі розвитку суспільства інформаційні технології стають головною причиною значного зростання актуальності сфер наукової діяльності. Процес моделювання реальних об'єктів дійсності переважно має на своєму шляху безліч труднощів, які з'являються вже під час формування завдань. Водночас створення комп'ютерних інтерфейсів, систем прийняття швидких рішень та автоматичного контролю в системах відеонагляду, які ґрунтуються на розпізнаванні мультимедійної інформації, стає головним завданням в процесі розвитку сучасних спеціалізованих та прикладних комп'ютерних програм.

Варто відмітити, що останнім часом розпізнавання образів користується все більшою популярністю, адже розпізнавання зображення, тексту чи мови, а також різноманітних явищ сприяє спрощенню комунікативного зв'язку людини з комп'ютером, допомагає застосовувати різні системи штучного інтелекту, в тому числі в системах відеоспостереження. Можливість сприймати зовнішній світ у формі образів сприяє передумовам дослідження властивостей величезної кількості об'єктів завдяки ознайомленню з кінцевою їх кількістю, а об'єктивна ознака властивостей образів допомагає створювати модель їх розпізнавання. Крім того, системи відеоспостереження, що використовують штучний інтелект, на сьогоднішній день активно розвиваються та з успіхом починають застосовуватись на масштабному рівні. Покращується як наукова база штучних нейронних мереж, так і обчислювальні потужності технічного обладнання.

Таким чином, постає необхідність у застосуванні сучасного підходу до проектування автоматизованих процесів систем охоронного призначення та контролінгу за допомогою використання комп'ютерного зору, а це, в свою чергу, вимагає залучення новітніх методів обробки рухомих об'єктів. Для

цього використовуються методи, які засновані на розпізнаванні образів, що сприяють вирішенню таких завдань:

- завдання ідентифікації;
- завдання кластеризації;
- завдання класифікації процесів, сигналів та різноманітних явищ.

Як відомо, використання методу розпізнавання рухів об'єктів в системі відеоспостереження, підвищує ефективність розробки автоматизованого процесу, зменшує матеріальні та часові витрати, сприяє отриманню об'єктивних та оперативних даних в режимі он-лайн.

Незважаючи на велику кількість вже існуючих систем розпізнавання образів, всі вони входять до складу більш великих систем, а тому для вирішення більш конкретних задач, наприклад, під час відслідковування рухів об'єктів в складі системи відеонагляду, дані системи не підходять.

Виходячи з вищенаведеного, проектування та практична реалізація програми розпізнавання рухів об'єктів за допомогою нейронної мережі для системи відеоспостереження є актуальною задачею.

Мета роботи – аналіз особливостей розробки автоматизованої моделі розпізнавання рухів об'єкту для системи відеонагляду.

Для досягнення мети були поставлені наступні *завдання*:

1. проаналізувати теоретичні засади дослідження системи розпізнавання рухів об'єкту;
2. дослідити алгоритм створення програми розпізнавання рухів об'єктів для системи відеонагляду;
3. спроектувати модель розпізнавання рухів об'єктів для системи відеонагляду;
4. розробити методичні рекомендації щодо використання методу розпізнавання рухів об'єктів в системі відеоспостереження.

Об'єкт дослідження – процес проектування програми розпізнавання рухів об'єктів для системи відеонагляду.

Предмет дослідження – специфіка застосування алгоритму виявлення рухомих об'єктів на зображенні в системах відеонагляду.

Методи дослідження:

- 1) методи системного аналізу;
- 2) аналіз існуючих методів;
- 3) спостереження;
- 4) абстрагування;
- 5) узагальнення.

Теоретично - інформаційну базу дослідження склали праці таких науковців, як Г. Аваліані, В. Антонюк, С. Вислоух, Г. Вороновський, Р. Вудс, Є.В.Главач, В. Глушков, У. Гренандер, А. Дейч, І. Дремін, В. Круглов, Е. Патрік, Р. Каллан, Д. Рутковська, Д. Тархов, Ф. Уосерман, Я. Фомін, Л. Шапіро, Л. Ясницький та інших.

Наукова новизна одержаних результатів. Результати дослідження пропонують альтернативний метод використання камер відеоспостереження, що базуються на штучному інтелекті.

Практичне значення одержаних результатів полягає в тому, що дослідження ґрунтується на результатах поглибленого вивчення особливостей застосування системи розпізнавання образів, а також розроблено модуль розпізнавання рухів об'єкту для систем відеонагляду.

Структура роботи. Дипломна робота складається зі вступу, трьох розділів, висновків, переліку посилань, додатків. Зміст роботи викладено на 65 сторінках. Перелік посилань складається із 31 найменування.

1 АКТУАЛЬНІСТЬ ЗАДАЧІ ВИЯВЛЕННЯ РУХОМИХ ОБ'ЄКТІВ

1.1. Актуальність теми

В наш час намітилася тенденція, пов'язана з широким впровадженням систем відеонагляду в різні сфери життя людини. Значний прогрес в цій сфері обумовлений появою досконаліших пристроїв отримання зображень, обчислювальної техніки, а також комунікаційних каналів. Найактуальнішими завданнями, які може вирішити система розпізнавання рухомих об'єктів, це:

- виявлення несанкціонованого проникнення на стратегічні об'єкти;
- контроль пересування людини в небезпечних зонах;
- виявлення несанкціонованого проникнення на стратегічні об'єкти;
- виявлення та спостереження в заданій області об'єктів інтересу або сторонніх об'єктів;
- фіксація порушень ПДР та небезпечних випадків на дорогах;
- виявлення залишених або забутих предметів;
- підрахунок числа людей (пасажирів, відвідувачів тощо);
- підрахунок в реальному часі щільності автомобільного потоку на трасах для кращого управління транспортною інфраструктурою;
- визначення кількості місць для паркування авто;

1.2. Проблеми задачі виявлення рухомих об'єктів

Зір – це головне джерело інформації про навколишній світ для людини. Тож технічна реалізація зорових здібностей – один з найважливіших кроків для того, щоб створити інтелектуальні системи, що будуть здатні забезпечити прогрес суспільства в багатьох різних напрямках. До таких напрямків відносяться:

- виготовлення принципово нових товарів;
- охорона навколишнього середовища (флори і фауни);
- значне скорочення злочинності;

- запобігання тероризму;
- запобігання катастроф та техногенних аварій, а також нещасних випадків в цілому;
- значне підвищення якості та зниження собівартості для промислового виробництва;
- організація та управління найважливішими процесами в багатьох сферах життя людей.

Однією з найактуальніших проблем, які на сьогоднішній день ще не отримали рішення – це аналіз послідовностей зображень динамічних сцен для оцінки руху об'єктів інтересу. Це завдання є дуже актуальним в таких областях, як:

- організація руху автомобільного транспорту;
- аналіз поведінки людей в місцях масового скупчення;
- боротьба з автомобільними пробками;
- оцінка ефективності реклами;
- оцінка загрози і запобігання терористичних актів;
- охорона території;
- контроль на виробництві.

Але не зважаючи на те, що завдання виявлення, оцінки параметрів, а також виявлення об'єктів загалом не є новим напрямком, наукове співтовариство та вчені всього світу продовжують вести активну роботу для усунення недоліків в роботі інтелектуальних систем відеоспостереження. Дуже багато статей і доповідей присвячено розвитку теорії і методів виявлення, оцінки параметрів та спостереження об'єктів, як додаток до завдання аналізу і обробки відеоданих. В багатьох різних сферах діяльності, де від людини необхідні механічні дії, розвинені інформаційні технології і сучасна робототехніка вже цілком успішно замінюють людину, при цьому, підвищуючи якість роботи та її надійність. Однак, далеко не завжди можливо замінити людину машиною. Це обумовлено тим, що в більшості випадків завдання розпізнавання, оцінки параметрів і виявлення об'єктів на

зображенні, існуючі на сьогоднішній день, як і раніше можуть бути вирішені тільки людиною в найбільш загальному вигляді.

1.3. Сфери використання та можливості систем розпізнавання

Теорія розпізнавання образів – це один із розділів інформатики, він базується на моделюванні основ та способів ідентифікації предмету та явища, або різноманітних сигналів. Необхідність розпізнавання образів виникає в багатьох галузях: від машинного зору, символічних розпізнавань, розпізнавання мови, і закінчуючи вузько спеціалізованими завданнями, таких як приладобудування[1].

Образ являє собою певну структуру в системі класифікації, яка акцентує конкретну групу об'єктів за вибраними ознаками. Образи мають характерні властивості, які проявляються в тому, що знайомство з кінцевою кількістю явищ з одного числа множини дозволяє дізнаватися про будь-яку кількість його представників. Образи володіють об'єктивними властивостями в тому аспекті, що люди, які вчаться на різних матеріалах спостережень, переважним чином однаково та незалежно від інших поділяють одні й ті ж об'єкти за певними критеріями оцінювання. На практиці в задачі розпізнавального процесу універсальна множина ділиться на певні частини-образи. При цьому відображення будь-якого об'єкта органами сприйняття системою розпізнавання незалежно від знаходження цього об'єкту щодо цих органів, називають зображенням об'єкта, а множини цих відображень, що об'єднані схожими якостями, називають образами [2].

Розпізнавання образів з англійської мови - «pattern recognition» - це теоретичний розділ науки штучного інтелекту, цей розділ вивчає способи класифікації об'єктів. Зазвичай об'єкт, який класифікується, являє собою образ (pattern). Образом може легко стати цифрова фотографія, наприклад, розпізнавання конкретних зображень; цифра чи літера, наприклад, у розпізнаванні символу; запис мови під час розпізнавання мови. Розпізнавання образів – задача визначення об'єкта або ідентифікація певних

якостей цього об'єкта за його зображенням, так зване оптичне розпізнавання, або за допомогою аудіо-запису в акустичному розпізнаванні, також за іншими властивостями. Водночас, образом є сукупність в системі класифікації, яка акцентує (з'єднує) конкретну групу об'єктів за певною ознакою.

Варто зазначити, що розпізнавання образів (рухів, символів) – це найфундаментальніша проблема інтелектуальних систем як науки. З іншого боку, завдання розпізнавання образів відіграє велику практичну роль [3].

Для вирішення задач розробки як автоматичної, так і автоматизованої системи керування теорія розпізнавання образів в якості зразка досить часто застосовує різні технології інтелектуальної діяльності суспільства. Наука у вигляді певного обладнання чи алгоритмів широко використовує можливості людського інтелекту. На думку дослідників, такий стан справ під час вирішення непростих та надскладних задач математичного характеру є найбільш раціональний, оскільки протягом еволюції людина навчилася удосконалювати власні можливості під час вирішення задач автоматичного керування. Око-процесор являє собою інтелектуальну систему, що спроможна створювати образне відтворення сучасної дійсності завдяки тому, що візуальні інформаційні потоки вона може сприймати в довільній формі. Також око-процесор здатний розрізняти та виокремлювати властивості явища або предмету, ознаки середовища. Він намагається обробити їх та видає рішення автоматично. Іноді рішення приймається з участю оператора. Система виконує такі дії [4]:

- попередня фільтрація;
передбачається усунення шумів та слабо зв'язаних точок зображення, виділення вихідного зображення;
- переміщення зображення;
під час даного процесу відбувається одночасне переміщення цілої картини на задану відстань по периметру цілої картини;
- зміна розміру зображення;

відбувається розширення картинки на відповідну величину, а також його ущільнення;

- повороти картинки;
- з'ясування центру картинки.;

під центром мають на увазі центр мас. Де під одиницею маси мається на увазі значення ij . Відповідні точки в площині властивостей, що задаються;

- виконання логічних операцій над зображеннями;
- виокремлення частини зображення.

Важливою рисою око-процесора є здатність приймати адаптивні рішення. Ця особливість зумовлюється наступними концептами[5]:

1. величезна кількість простору властивостей;
2. присутність алгоритмів прийняття рішень, які несуть в собі вагові показники;
3. наявність малопередбачуваних ситуацій в умовах реального функціонування;
4. присутність обставин з пороговим значенням прийняття рішень;
5. експлуатація за обставин апріорного невизначення.

Отже, метод розпізнавання образів користується популярністю, в тому числі в галузі штучного інтелекту. Варто сказати, що явище розпізнавання образів ґрунтується в переважній більшості на певній подібності об'єктів. При цьому, як відомо, більшість предметів мало схожі один до одного. Проте, як показує практика, серед деяких об'єктів завжди знайдеться певна схожість за тією або іншою властивістю. При цьому методи розпізнавання образів поділяють на два види[6]:

1. структурні;
2. засновані на теорії рішень.

Як правило, теорію рішень використовують при обчисленні за допомогою величини кількісного характеру. Структурні орієнтуються на образи, опис яких передбачає якісніші характеристики.

В процесі розпізнавання образу велику роль має навчальна вибірка, іншими словами, так звана впорядкована сукупність певних властивостей. Класом образів прийнято називати сукупність об'єктів з подібними ознаками. Здебільшого користуються такими формами властивостей, як властивість дерева, векторна ознака та ознака символного рядка.

Способи, що ґрунтуються на зіставленні, являють собою компоненти векторних властивостей в кожному класі об'єктів. Створений образ має відноситися до того класу, що буде найближчим у площині заздалегідь заданої метричності. Логічно, що простий підхід – це пошук мінімальної відстані, яка обчислюється завдяки евклідовим нормам між конкретним набором векторів властивостей невідомого об'єкта та набором векторів майбутньої моделі. Висновок про те, чи належить об'єкт до певного класу, проходить за найменшою відстанню. Мінімальний розподіл відстані функціонує краще за тих обставин, коли відстань між точками математичних очікувань класів більша за діапазоном розкидів об'єкту по кожному окремо взятому класу.

З іншого боку, велике значення мають способи розпізнавання образів, які ґрунтуються на імовірнісних показниках за допомогою випадковостей, які впливають на моделювання нових класів образів [7].

Таким чином, потрібно виробити оптимальний алгоритм, який би значно зменшив прогнозовану кількість ймовірних помилок. Водночас варто зауважити, що досить не простим завданням є можливість зрозуміти, як саме має виглядати оптимальний спосіб з точки зору опису комп'ютеру. Так, відомо, що всі існуючі способи поділяють на три ступені:

- процес первинної обробки;
- процес відсіювання та логічне з'ясування результатів фільтрації;
- процес прийняття рішення.

Зазвичай, для науки розпізнавання образів для того, щоб розпізнати предмет чи явище на малюнку, - потрібно використати відразу всі вище зазначені групові процеси. Проте іноді буває достатньо двох. До етапу

фільтрації відносяться способи, які допомагають ідентифікувати на зображенні ті об'єкти, що зацікавили в даний момент, без їхнього попереднього дослідження. Більшість таких способів використовує єдиний алгоритм на всі точки зображення одночасно. На цьому рівні аналізування, як правило, не використовується [8].

Найпростішим перетворенням є бінаризація зображення по порогу. Для зображень відтінків сірого кольору, наприклад, такою межею є значення яскравості. Визначення та обрання порогу бінаризації дозволяє визначити вид самого алгоритму. Зазвичай бінаризація відбувається під час процесу адитивного визначення порогів. Так, такою операцією може бути вибір математичного очікування.

Варто додати, що класичні способи фільтрації використовуються в широкому спектрі задач. Найпоширенішим з класичних способів перетворення є трансформація Фур'є. Однак даний спосіб не застосовується в зображеннях як таких. Проте для дослідження зображень іноді не вистачає простого перетворення одновимірності та виникає необхідність у ресурсномістких двовимірних перетвореннях [8].

Обчислення за формулами є досить трудомістким, тому практично використовують згортку потрібного участку з використанням низькочастотної або високочастотної фільтрації, в залежності від конкретної задачі. Слід відмітити, що такий процес спрощень не дає досягти широкого спектру операцій на кшталт аналізу, проте часто достатнім є сам результат без подальшого перетворення.

Перетворення Вейвлет є більш перспективним способом аналізу зображень на відміну від способу Фур'є. Метод спрощує процес компонування, дослідження та передачі великого об'єму образів. Вейвлет-трансформації ґрунтуються на тому, що розділяють зображення на малі частотні хвилі, при цьому частота хвиль постійно рухається в часовому відрізьку часу у порівнянні зі способом Фур'є.[9].

З математичного ракурсу образи картинки є двовимірною матрицею відповідних заданих величин яскравості. Проте під час переходу від однієї його частини до іншої гістограми доволі часто суттєво змінюються. Прикладом використання вейвлет-перетворення є задача пошуку відблискуочей, де вейвлето - це, власне, відблиск. В основі цього методу-кореляційність, що застосовується переважним чином разом з іншими способами, проте може використовуватися і окремо. В процесі розпізнання образу в картинці кореляційність відіграє основну роль.

Наступним популярним класом фільтрації є фільтрація функцій. Фільтрація дозволяє на відображенні знайти шматочки простих функцій. Найбільш відомим є перетворення Хафа, яке допомагає знайти різноманітні ефективні обчислювальні функції. Схожим способом є метод перетворення Радону, який завдяки обчисленню крізь швидке перетворення Фур'є виграє в оптимізації. Ще один розділ фільтрації - це фільтрація різноманітного контуру. Така фільтрація є корисною в ситуації, якщо об'єкт достатньо складний, проте існують чіткі границі. Тоді фільтрація контурів є однією з засадничих механізмів роботи із зображеннями. Після фільтрації виходить набір даних, що потребують обробки. Проте іноді вони потребують додаткового логічного перетворення. Через те виникає необхідність у методах, які дозволяють переходити від цілого зображення до конкретних ознак об'єктів на ньому [10].

Також існують методи, які допомагають визначати об'єкт по контурам. Такий метод називається контурним аналізом. Обрані точки являють собою визначальні індивідуалістичні властивості, що допомагають зіставити різні класи об'єктів.

Способи машинного навчання та прийняття рішень - це кінцева стадія в процесі розпізнавання. Ці способи знаходяться на межі статистичної математики, способів оптимізації, а також класичних наук математики, але мають також і власну особливість, яка пов'язана з проблематикою обчислювальної продуктивності та процесом перенавчання. Здебільшого суть

навчання полягає в наступному: на базі певної навчальної вибірки з властивостями кожного класу необхідно спроектувати таку систему, завдяки якій комп'ютер зможе дати аналіз новому зображенню та визначити, який з об'єктів представлений на зображенні [11].

Варто зазначити, що людський зір без проблем адаптований до сприйняття оточуючого середовища. Наприклад, він може відокремити головний об'єкт від його оточення, врахувати тонкі закономірності затінення (Рис. 1.1.)



Рисунок 1.1 - Приклад людського зору

Водночас дивлячись на фотографію групи осіб (Рис 1.2), можна без проблем визначити точне число людей, а також припустити, в якому вони зараз настрої відповідно до їхнього виразу обличчя. Психологи, які займаються вивченням процесів сприйняття, витрачають не один десяток років аби зрозуміти процес функціонування візуальної системи. Навіть використовуючи під час досліджень спеціальні оптичні ілюзії для знаходження деяких фундаментальних засад та об'єктивних рішень в цій площині наразі немає [12].



Рисунок 1.2 - Колективне фото

Слід додати, що водночас дослідники комп'ютерного зору розвивали галузь математичних моделей розпізнавання образів (символів, рухів). Наприклад, за допомогою доволі великого набору даних, можливо проектувати точні, щільні 3D- моделі, з використанням принципів стереопідсилювання (Рис. 1.3)

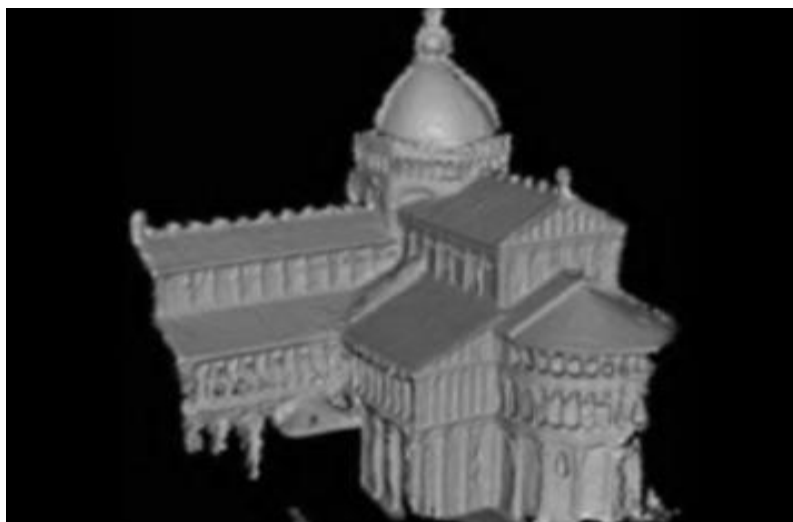


Рисунок 1.3 - Приклад 3D- моделі

На сьогоднішній день сучасні алгоритми розпізнавання руху об'єкту спроможні відстежити переміщення особи чи групи осіб за доволі складних технічних умов зйомки (Рис. 1.4)



Рисунок 1.4 - Приклад розпізнання руху об'єкту

Проте, незважаючи на великі досягнення, ступінь розпізнавання рухів (символів, образів) та виокремлення контексту із конкретного зображення досі перебуває на низькому рівні розвитку. Причиною є проблема зворотної природи, де потрібно відновити певну невідому інформацію на базі невеликої кількості даних.

Проте моделювання візуального світу в реальному часі з усім його розмаїттям неможливий на даному етапі функціонування світу[13].

У алгоритмах комп'ютерного відображення дійсності відбуваються спроби зворотного відновлення даних, які приховуються, іншими словами, спроба описати світ, що представлений в одному або серії образів та відтворити його якості на кшталт форми, освітленості та переходів кольорів. Люди, які не працювали у цій галузі, часто недооцінюють труднощі проблеми.

Це неправильне розуміння того, що бачення повинно бути легким з появою штучного інтелекту, коли вважали, що когнітивні компоненти штучного мозку були по суті складнішими, ніж перцептивні елементи.

Таким чином, спроможність сприймати зовнішній світ у вигляді образів дозволяє з конкретною ймовірністю досліджувати характеристики

безлічі об'єктів завдяки ознайомленню з кінцевим їх числом, а характерні ознаки образів допомагають в процесі їх розпізнавання.

Для того, щоб краще розібратися в тематиці введемо наступні поняття:

Відеокадр - растрове зображення у форматі RGB.

Потік відеокадрів - послідовність відеокадрів, отримана від цифрового пристрою реєстрації відеозображень (цифрової відеокамери).

Рух - поступальний рух деякої групи пікселів, яке спостерігається на сусідніх кадрах відео потоку. необхідною умовою руху є можливість локалізації однієї і тієї ж групи пікселів на сусідніх кадрах.

Рухомий об'єкт - це група пікселів, яка зберігає свою форму на сусідніх кадрах, а обрамляє її прямокутник зберігає розміри, причому його центральна точка відчуває мале зміщення між цими кадрами (в межах декількох пікселів). На рисунку 1.5 представлена група пікселів в прямокутнику. Центр групи розташований в точці $C_i (x_i, y_i)$ і $C_{i+1} (x_{i+1}, y_{i+1})$ на i -му та $i+1$ -му кадрі, відповідно. Зауважимо, що форма групи пікселів може змінюватися, але це зміна зазвичай мізерно, і їм можна знехтувати, вважаючи, що форма залишається незмінною.

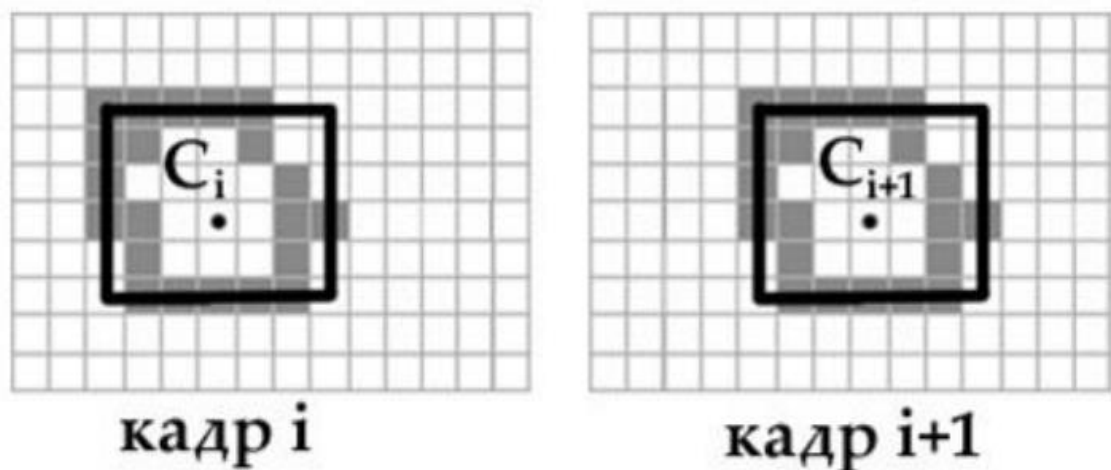


Рисунок 1.5 - Група пікселів на сусідніх кадрах

Відео-дані - це потік даних, що мають структуру послідовних кадрів відеозображення. Дані кадри можуть бути спочатку отримані від приладу реєстрації (цифрової / аналогової камери) або синтезовані за допомогою спеціальних програм, а потім неодноразово перетворені для зручності їх

збереження і передачі. Всі кадри подібного потоку мають однаковий розмір зображення і надходять з певним незмінним тимчасовим інтервалом.

Величина, зворотна часового інтервалу між сусідніми кадрами, - це число кадрів, яке потрібно обробити за секунду (frames per second або fps). Крім цього, всі кадри потоку мають схожу глибину кольору, яку можна виміряти як кількість використовуваних бітів пам'яті для зберігання 1 пікселя кадру (bits per pixel або bpp). У дипломній роботі дослідження будуть вестися тільки з цифровими відеоданими, іншими словами, з тими, які зберігаються в пам'яті ЕОМ і обробляються на ній програмно або апаратно.

1.4. Передобробка і постобробка зображень при виявленні рухомих об'єктів

Алгоритми в просторовій області базуються або на застосуванні інтенсивності пікселів зображень, або на виділення ознак: колір, статистики, текстури і так далі. У загальному випадку, алгоритм виявлення рухомого об'єкту на відеопослідовності зводиться до чотирьох основних етапів:

- попередня обробка (передобробка),
- моделювання фону,
- виявлення областей зображень, відповідних рухомим об'єктам (виділення переднього плану),
- постобробка.

1.4.1. Попередня обробка (передобробка)

Метою методів виявлення рухомих об'єктів на відео-даних є виявлення значущих змін, які з'являються в відеопослідовності, при ігноруванні малозначних. Дії по передобробці зображень включають геометричні коригування та коригування яскравості [3]. Також може застосовуватися інформація про відстань до об'єкта (глибині).

Коригування яскравості.

Існує кілька методів компенсації зміни освітленості між кадрами, які виникають через зміни яскравості або позиції джерела світла сцени. Один з таких методів: нормалізація значень яскравості пікселів зображення. Він застосовується для отримання конкретного значення математичного очікування і дисперсії [4, 5].

За даним методом можуть бути нормалізовані як поточний кадр, так і фонове зображення (отримання нульового математичного очікування з одиничної дисперсією) [3]. Однак такий підхід дозволяє отримати кращі результати для локальних змін освітленості в кадрі ціною появи додаткових артефактів [3]. Алгоритми, що використовують нормалізацію, в основному, дають погані результати для темних областей зображення.

Інформація про глибину (відстані до об'єктів).

Завдяки застосуванню цих даних виникає менше проблем при виявленні об'єктів, які рухаються, так як вони в найменшій мірі схильні до спотворення в зв'язку зі змінами умов освітленості. Але методи, які залежать від відстані, є неефективними при використанні їх для малоконтрастних областей або областей, які недоступні для всіх камер. Даний метод також часто не справляється з виділенням переднього плану (рухомих об'єктів), що знаходяться в безпосередній близькості від фону, наприклад, рука на стіні, ноги на підлозі і т.д.

1.4.2 Постобробка

Поділ об'єктів.

Поділ виявлених і виділених об'єктів, що знаходяться на достатній відстані, не представляється складним завданням. В даному випадку рішенням є використання простого порога. Важким завданням є поділ пересічних або знаходяться поруч об'єктів [6]. Використання певних методів обумовлено конкретною ситуацією і необхідною точністю виділення об'єкта.

Існують підходи поділу рухомих об'єктів, засновані на побудові мережі (дерева графів), як в завантаженому відео, так і в режимі реального часу [7, 8].

Виділення меж (контуру).

Точність виділення контуру об'єкта, що рухається залежить від точності виділення області рухомого об'єкту. Виділяють такі помилки виділення області рухомого об'єкту, як шум, помилкове виявлення (помилка першого роду) і пропуск цілі (помилка другого роду). Для вирішення даних проблем в основному застосовують метод відповідної фільтрації на стадії виявлення [9].

Головними причинами появи шумів, які надають неминуче вплив на результат виділення об'єкта, є зміна освітленості, зміни положення камери, колірний шум, шум системи. Для усунення шуму використовується адаптивна фільтрація. Також застосовується сигма-дельта фільтр на стадії віднімання фону для отримання більш точної маски об'єкта. Цей метод дозволяє зменшити вплив нестабільності фону і шумів, викликаних невеликим рухом камери або фону.

1.4.3 Математична морфологія

Математичної морфологією вважають аналіз зображення з точки зору його форми. При використанні даного підходу над зображеннями виконують ряд перетворень змінюють форму об'єктів, що містяться на зображенні [10].

Математичну морфологію використовують в різних системах, які працюють над обробкою зображень, на різних етапах і для досягнення різних цілей:

- Покращення візуальних характеристик зображення (яскравість, контрастність і т.д.)
- Відновлення зіпсованих зображень, наприклад, реставрація фотознімків.
- Виявлення контурів.
- Зниження рівня шуму.

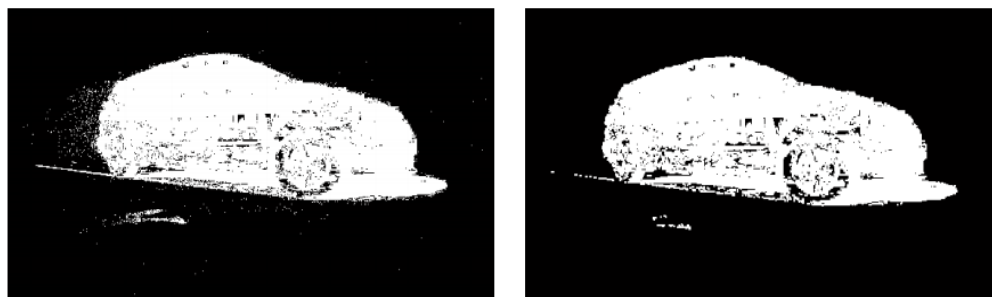
Метод математичної морфології може застосовуватися по відношенню до

кольоровим, чорно-білим зображенням і до зображень у відтінках сірого. Для цих трьох випадків формально даний підхід визначається кілька по-різному. Розглянемо випадок бінарних чорно-білих растрів. Так як при використанні операцій математичної морфології над бінарними чорно-білими растрами також можна провести додаткову фільтрацію шуму в масці руху. Цей процес вважається ефективним, оскільки маска складається з нулів і одиниць. Існує велика кількість операцій математичної морфології: перенесення, нарощування, ерозія, закриття, а також відкриття [11].

Відкриття (Opening) двійкового B на структурний елемент S записується як $B \circ S$ і визначається як:

$$B \circ S = (B - S) \oplus S$$

Приклад роботи операції відкриття представлений на малюнку 1.6, де наведені вихідна маска руху до і після обробки.



а)

б)

Рисунок 1.6 - Приклад роботи операції відкриття

а) Маска руху до обробки

б) Маска руху після обробки

За представленими вище малюнків можна сказати, що операція відкриття видаляє з зображення всі елементи, розмір яких менше розміру структуруючого елемента. Інші елементи зображення зберігаються.

Висновки за розділом 1

Підсумовуючи перший розділ, можна зробити наступні висновки:

1. Визначили теоретичні засади методів розпізнавання рухів об'єкта;
2. Визначили проблеми задачі виявлення рухомих об'єктів;
3. Охарактеризували класифікацію системи розпізнавання рухів об'єкту;
4. Розглянули задачі розпізнавання образів;
5. Розглянули сфери використання систем розпізнавання.
6. Розглянули поняття переобробки, постобробки зображень та математичної морфології.

2 МЕТОДИ ТА АЛГОРИТМИ РОЗПІЗНАВАННЯ РУХОМИХ ОБ'ЄКТІВ

2.1. Огляд методів виявлення рухомих об'єктів

Завдання виявлення рухомих об'єктів була поставлена досить давно, хоча внаслідок своєї специфіки, вона поки не має конкретного рішення.

По-перше, умови виявлення бувають різні. Наприклад, якщо працювати з бінарним чорно-білим зображенням, то завдання значно спрощується в порівнянні з використанням повнокольорових растрів з плавними переходами кольору від одного відтінку до іншого. Відео може надходити з статично розташованої камери, таким чином, всі кадри будуть мати приблизно однаковий фон з можливими відмінностями освітлення.

Також камера, розміщена на об'єкті, що рухається, може знімати нерухомий об'єкт. Рівень шуму може значно змінюватися. Природні явища: дощ, туман, вітер, сніг можуть бути причиною істотних коливань в спочатку нерухомої сцені. Все це призводить до тому, що методи, неперевірено працюють в одних умовах, зовсім непридатними для інших.

Алгоритми, що працюють в автоматичних системах виявлення руху, - особливо це стосується систем контролю периметра, інших систем виявлення проникнення - повинні бути стійкі для роботи в широкому спектрі значно різняться зовнішніх умов. Нижче представлені деякі вимоги до подібних методів:

- Невисока обчислювальна складність. Це можливо найважливіша характеристика, яка оцінюється в момент вибору детектора руху. Її висока значимість обумовлена тим, що при великому обсязі обчислень, які буде проводити детектор, подібні методи неможливо буде застосовувати в режимі реального часу.
- Стабільна робота в будь-який час року при різних погодних умовах.
- Сталий виявлення руху в різний час доби при штучному освітленні.

- Мінімум помилкових спрацьовувань на шум реєструє і передавальної апаратури.
- Мінімальна кількість спрацьовувань на об'єкти, що не мають будь-якого значення (хитні дерева, листя, інша рослинність тощо).

Вимоги до систем виявлення рухомих об'єктів, крім перерахованих вище, що стосуються алгоритмів виявлення руху, зазвичай виглядають так:

- Багатоканальність.
- Масштабованість. Система повинна легко збільшувати кількість джерел вхідних даних.
- Можливість гарячої заміни. Необхідно мати можливість підключення / відключення будь-якої камери, без припинення на якийсь час роботи інших каналів.

Першим кроком у виявленні об'єктів, які рухаються є знаходження всіх його можливих варіантів на зображенні. В основному це робиться за допомогою методів, які засновані на моделюванні. Вони порівнюють окремі області зображення з отриманими моделями рухомого об'єкту.

Існують різні підходи для вирішення завдань по виявленню рухомого об'єкту. Вони можуть бути розділені на 2 класи: методи, базуються на тимчасових відмінностях (метод обчислення оптичного потоку), і моделювання і віднімання заднього фону.

Перші наукові праці, засновані на виявленні рухомих об'єктів, написані в кінці 70-х років ХХ століття, базувалися на відмінності сусідніх відеокадрів. Головними правилами всіх алгоритмів виявлення об'єкта 70-80-х років були так звані аспекти статистичних гіпотез. У 90-х роках ХХ століття в основному розвивалися методи, засновані на зміні положення пікселів зображення. Застосовувалися як будь-який окремий піксель, так і блоки пікселів. У 90-х роках ХХ століття був запропонований комплексний метод пізнання [12], ну а в 2005 році [13] запропонований метод «Від протилежного», що базується на принципі Гельмгольца. Метою цього методу було, власне розпізнавання відеокадру, що не має рухомого об'єкта.

Методи, які базуються на відмінності сусідніх кадрів, надзвичайно чутливі до шуму, а також до змін освітленості. Якщо кількість кадрів в послідовності велике, а також між сусідніми кадрами є деякі невеликі зміни, іншим рішенням для виявлення рухомого об'єкту вважається моделювання і віднімання фону. В даному випадку пікселями переднього плану будуть пікселі, для яких значення інтенсивності відмінно від середнього значення моделі фону.

Для виявлення змін на передньому плані шляхом вирахування фону використовується якась модель. Дана модель фону оновлюється через певний інтервал часу. В якості найпростішої моделі іноді застосовується перший відеокادر [14]. При відніманні цього зображення з подальших рухомих об'єкт представляється досить грубо. Дана схема віднімання може використовуватися виключно з простими сценами з однорідним контрастом і незначними змінами.

Існують різні методи і алгоритми, які описують рішення задач по виділенню фону за допомогою використання статистичних функцій: функції середнього, медіани, а також за допомогою розподілу Гауса. Даний метод моделює кожну точку зображення за допомогою розподілу Гауса з обчисленим середнім значенням інтенсивності.

На рисунку 2.1 представлений приклад виявлення об'єктів, що рухаються.

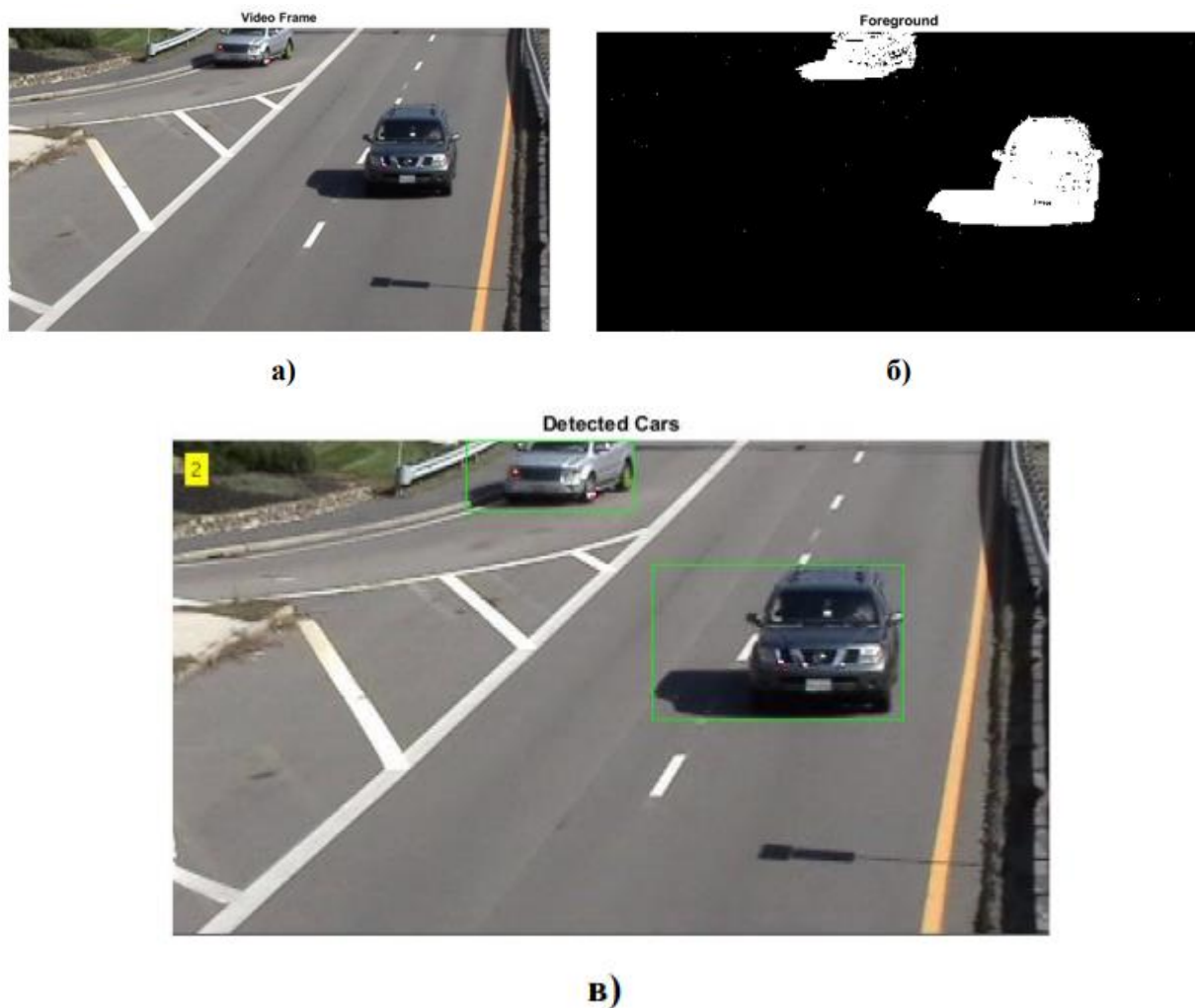


Рисунок 1.6 - Приклад виявлення рухомого об'єкту

а) Початкове зображення

б) Бінарне зображення

в) Виділення рухомих об'єктів

Також існують способи виявлення об'єктів з використанням частотної області, які базуються на вейвлет-перетвореннях. В деяких випадках вейвлет-перетворення застосовується для оцінки фону на базі попередніх кадрів відеопослідовності, також для виділення пересувається об'єкта і його розташування [15]. Існує інший спосіб, базується на дробовій статистикою низького порядку [16].

2.2. Метод віднімання фону

Простим методом до вирішення завдання виділення рухомого об'єкту є використання механізму віднімання фону з кадру відео (background subtraction) [17]. Процедура віднімання передбачає, що для даного відео побудована модель фону

$$F = \{F(x, y), 0 \leq x < \text{width}, 0 \leq y < \text{height}\},$$

Де width - ширина кадру, height - висота кадру.

В принципі, така модель повинна являти собою сцену без рухомих об'єктів. При цьому необхідно її регулярно оновлення, для того щоб враховувати зміну умов освітленості і налаштувань камери, таких як поворот, нахил і зміна фокусної відстані. Іноді використовують процедуру поновлення моделі фону з плином часу.

Процедура вирахування моделі фону складається з двох етапів:

- Віднімання фонового зображення з поточного кадру відео. На даному етапі відбувається попіксельне віднімання інтенсивності кадру відео і фонового зображення:

$$D_k(x, y) = \text{abs}(I_k(x, y) - F(x, y)), k = \overline{1, N},$$

де $I_k(x, y)$ – інтенсивність певного кадру.

- Поділ пікселів за належністю до фону і об'єкта, тобто побудова бінарного зображення (маски). Вважається, що піксель належить об'єкту і має білий колір в масці, якщо різниця інтенсивності фону і поточного кадру для даного пікселя перевищує деяке порогове значення, в іншому випадку, приймається, що піксель належить фону:

$$M_k(x, y) = \begin{cases} 255, & D_k(x, y) \geq \tau \\ 0, & D_k(x, y) < \tau \end{cases}, k = \overline{1, N}$$

Від якості побудованої моделі фону залежить якість визначення положення рухомого об'єкту за допомогою методу віднімання фону. Усе

методи вирахування фону діляться на дві групи в залежності від механізму побудови фонового зображення:

1) Нерекурсивні методи.

Нерекурсивні методи при оновленні моделі фону для поточного кадру використовують дані про інтенсивності пікселів деякого набору попередніх моделей фону [18] (або кадрів) і поточного кадру. Деякі методи, які відносяться до нерекурсивних, представлені нижче:

- Метод віднімання поточного і попереднього кадру. В даному методі передбачається, що для певного кадру I_k модель фону F_k

збігається з попереднім кадром, тобто $F_k = I_{k-1}$. З цього випливає,

що на першому етапі алгоритму віднімання фону обчислюється різниця двох послідовно йдуть кадрів відеопослідовності:

$$D_k(x, y) = \text{abs}(I_k(x, y) - F_k(x, y)) = \text{abs}(I_k(x, y) - I_{k-1}(x, y)), k = \overline{2, N}$$

- Метод усереднення деякої кількості попередніх кадрів. В даному випадку кількість кадрів, за якими виконується побудова моделі фону, позначається як S . Тоді модель фону F_k для кадру I_k можна виразити наступною формулою:

$$F_k(x, y) = \frac{1}{S} \sum_{j=0}^{s-1} I_{k-j}(x, y)$$

- Метод визначення медіани фіксованої кількості попередніх кадрів. Також як в методі усереднення припускаю, що S - число кадрів, на підставі яких буде оновлюватися модель фону. Тоді модель фону F_k визначається за формулою:

$$F_k(x, y) = \text{med}_{j=0, s-1} \{I_{k-j}(x, y)\}$$

- Метод, заснований на використанні мінімаксного фільтра. За допомогою навчальної послідовності без рухомих об'єктів визначаються три значення для кожного пікселя фонового

зображення: мінімальна яскравість $z_{i,jmin}^0$, максимальна яскравість $z_{i,jmax}^0$, а також максимальна зміна яскравості між сусідніми кадрами $\Delta z_{i,j}^0$ [19].

Дані параметри обчислюються за формулами:

$$z_{i,jmin}^0 = \min z_{i,j}^0, z_{i,jmax}^0 = \max z_{i,j}^0, \Delta z_{i,j}^0 = \max |z_{i,j}^t - z_{i,j}^{t-1}|$$

Ці значення визначаються виходячи з параметрів декількох кадрів і періодично оновлюються для моделей фону.

Головною перевагою нерекурсивних методів є простота реалізації і швидкість оновлення моделей фону при переході від одного кадру до іншого в відеопослідовності. Однак точність роботи даних методів сильно залежить від швидкості об'єктів. Чим повільніше рухається об'єкт, тим гірше він визначається. Методи даної групи дуже неефективні при наявності рухомого фону (листя дерев, струмлива вода і так далі).

Для згладжування зазначених ефектів застосовують так зване α -зміщення. У даній процедурі оновлена модель фону для кадру I_k представляється опуклою оболонкою моделі фону F_{k-1} . Виражається наступною формулою:

$$F_k(x, y) = \alpha I_k(x, y) + (1 - \alpha) F_{k-1}(x, y)$$

2) Рекурсивні методи.

У даній категорії методів для поновлення моделі фону використовується дані про інтенсивності пікселів тільки поточного кадру. До рекурсивних методів можна віднести: гістограмний метод, метод представлення моделі фону сумішню гаусових розподілів (Gaussian mixture model) [20], метод "Шифрувальної" книги (codebook) [21], метод вилучення візуального фону (Visual Background Extractor, ViBe) [22]. Деякі методи представлені нижче.

- Гістограмного метод. В Відповідно до даного методу все колірний простір розбивається на окремі біти. У півтоновому зображенні як даного простору передбачається відрізок зміни інтенсивності, а в кольоровому зображенні - тривимірний куб.

Гістограма будується для всіх зображень в відеопослідовності.

Гістограмний метод здійснює аналіз всіх пікселів зображення. Виходячи з того, яка інтенсивність / колір спостерігається в пікселі, збільшується на одиницю значення відповідного біт гістограми. Якщо величина певного біта менше фіксованого порогового значення, то пікселі, що входять до складу цього біта, належать фону. В іншому випадку, приймається, що пікселі належать об'єкту. Головною проблемою використання гістограмного методу вважається необхідність використання додаткової пам'яті і виконання безлічі операцій звернення до пам'яті в процесі реалізації.

- Метод, заснований на використанні апроксимуючого медіанного фільтра. Підхід, де застосовувався простий рекурентний фільтр для оцінки медіани, був використаний для моніторингу міського транспорту [23]. В даному методі поточне значення медіани збільшується на одиницю, якщо піксель наступного кадру більше певної оцінки, і зменшується на одиницю в зворотному випадку:

$$z_{i,j}^{0(t)} = \begin{cases} z_{i,j}^{0(t-1)} + 1, & \text{якщо } z_{i,j}^t > z_{i,j}^{0(t-1)}, \\ z_{i,j}^{0(t-1)} - 1, & \text{якщо } z_{i,j}^t < z_{i,j}^{0(t-1)}, \\ z_{i,j}^{0(t-1)}, & \text{якщо } z_{i,j}^t = z_{i,j}^{0(t-1)}, \end{cases}$$

де $z_{i,j}^{0(t)}$ - опорне зображення для кадру t .

В результаті оцінка медіани зводиться до значення, коли половина пікселів, що входять стає більше цього значення, а половина - менше. Це і є медіаною. Недолік методу: повільна адаптація при досить великих змінах фону [24]. Для адаптації до області фону, яку покинув об'єкт, довгий час займав її, потрібно значну кількість кадрів.

- Суміш Гаусових розподілів (Gaussian Mixture Model - GMM) [20, 25]. При побудові фону за цим методом вважається, що для будь-якого пікселя (x_0, y_0) зображення I_k відома історія зміни його інтенсивності/кольору на всіх попередніх кадрах $\{X_1, X_2, \dots, X_k\} = \{I_j(x_0, y_0), j = \overline{1, k}\}$.

Тоді ймовірність спостереження значення X_k , представляється сумішшю гауссових розподілів:

$$P(X_k) = \sum_{j=1}^s \omega_j^k N(X_k | \mu_j^k, \Sigma_j^k),$$

де w_j^k - вага j -ого розподілу Гауса для кадру з номером k ,

μ_j^k - математичне очікування,

Σ_j^k - середньоквадратичне відхилення,

$N(X_k | \mu_j^k, \Sigma_j^k)$ - функція щільності нормального розподілу, яка виражається наступною формулою:

$$N(X_k | \mu_j^k, \Sigma_j^k) = \frac{1}{(2\pi)^{\frac{D}{2}} |\Sigma_j^k|^{\frac{1}{2}}} e^{-\frac{1}{2}(X_k - \mu_j^k)^T (\Sigma_j^k)^{-1} (X_k - \mu_j^k)}$$

Передбачається, що складові кольору є незалежними і мають однакову середньоквадратичне похибка. Виходячи з цього матрицю коваріації можна записати, як:

$$\Sigma_j^k = (\sigma_j^k)^2 E,$$

де E - одинична матриця.

Виходячи з цього твердження, можна зменшити обчислювальну трудомісткість методу за рахунок відсутності необхідності обчислювати матрицю, зворотну до матриці коваріації Σ_j^k . Таким чином, задано розподіл

спостережуваних значень кольору для кожного пікселя. Нове значення кольору можна уявити як одну з основних компонент отриманої суміші Гаусових розподілів. Також воно може використовуватися для поновлення параметрів моделі. Побудовані розподілу сортуються в порядку зменшення величини:

$$r_j^k = \frac{\omega_j^k}{\sigma_j^k}.$$

При даному сортуванні піксель фону відповідає розподілу з великою вагою і малою дисперсією.

Перші B^k розподілів можна віднести до кольору фонових пікселів, коли

$$B^k = \arg \min_b \left\{ \sum_{j=1}^b \omega_j^k > T \right\},$$

де T - порогове значення, параметр моделі.

При появі наступного кадру I_{k+1} для кожного пікселя зображення виконується тест, за допомогою якого при використанні відстані Махаланобіса, визначається, якому розподілу відповідає отримане значення.

$$\sqrt{(X_{k+1} - \mu_j^k)^T (\Sigma_j^k)^{-1} (X_{k+1} - \mu_j^k)} < 2.5 \sigma_j^k$$

Тоді можливі дві ситуації:

1. Якщо знайшовся відповідний розподіл Гаусса, то в залежності від того, чи визначає він розподіл фонових пікселів (входить в групу з B^k розподілів) чи ні, поточний піксель визначається як фоновий або як той, що належить об'єкту.

2. У разі, якщо жоден розподіл не відповідає розподілу Гаусса, що задовольняє умові, то вважається, що піксель належить об'єкту.

За даним правилом формується двовимірна маска.

Перед тим, як приступити до роботи над наступним кадром, потрібно оновити установки розподілів: математичне очікування та

середньоквадратичне відхилення. Оновлення виконується по-різному в залежності від результату знаходження відповідного розподілу для кольору поточного пікселя:

1. Відповідність виявлено. В такому випадку вагові коефіцієнти, які становлять суміш Гауссових розподілів, яким відповідають, і параметри розподілів перераховуються за такими формулами:

$$\omega_j^{k+1} = (1 - \alpha)\omega_j^k + \alpha,$$

$$\mu_j^{k+1} = (1 - \rho)\mu_j^k + \rho X_{k+1},$$

$$(\sigma_j^{k+1})^2 = (1 - \rho)(\sigma_j^k)^2 + \rho(X_{k+1} - \mu_j^{k+1})(X_{k+1} - \mu_j^{k+1})^T,$$

де α – задана константа,

$$\rho = \alpha N(X_{k+1} | \mu_j^k, \Sigma_j^k).$$

У разі, якщо для всіх розподілів X_{k+1} не відповідає, параметри не змінюються, однак перераховуються коефіцієнти w_j^{k+1} за такою формулою:

$$\omega_j^{k+1} = (1 - \alpha)\omega_j^k$$

2. Відповідність не виявлено. Тоді останнє (в сенсі введеного відношення порядку) розподіл Гаусса заміщається розподілом з новими параметрами. Математичне очікування стає рівним поточному значенню кольору пікселя:

$$\mu_s^{k+1} = X_{k+1}$$

Дисперсія $(\sigma_s^{k+1})^2$ вибирається максимально можливою, а вага w_s^{k+1} – мінімально допустимою.

Кількість розподілів вибирається виходячи з складності фону і наявної обчислювальної потужності.

- Метод вилучення візуального фону (Visual Background Extractor, ViBe) [22]. В даному методі модель фону на k -му кадрі

представляється набором множин $M^k(p) = \{v_1, v_2, \dots, v_N\}$ для всіх пікселів $p = (x, y)$, де v_i – інтенсивність/колір пікселя.

Для класифікації пікселя p в колірному просторі будується сфера $S_R(v(x))$, радіус якої R . Також обчислюється кількість векторів множини $M(p)$ яка потрапляє в побудовану сферу.

$$K = |S_R(v(p)) \cap M(p)|,$$

$$K > T_{\min},$$

де T_{\min} – фіксоване порогове значення.

Якщо виконується ця умова, то приймається, що піксель p належить фону, в іншому випадку, об'єкту.

На першому кроці повинна бути виконана ініціалізація множин $M^0(p)$ для всіх пікселів p за формулою:

$$M^0(p) = \{v^0(y), y \in N_G(p)\},$$

де $N_G(p)$ – околиця пікселя розміру 3×3 (9 клітин, включаючи поточний піксель);

у обирається N разів випадковим чином.

Оновлення моделі фону для кадру I_k складається з двох етапів:

1. Якщо p є пікселем фону, то з множини $M^k(p)$ випадково вибирається компонента, яка замінюється значенням $v(p)$.

2. Випадковим чином вибирається один сусідній піксель з околиці $N_G(p)$, для якого виконується попередній крок.

Існує безліч методів побудови моделей фону. У даному розділі представлені тільки деякі з них. В даний час ведуться активні дослідження в цьому напрямку.

2.3. Метод оптичного потоку

Оптичний потік являє собою зображення видимого руху об'єкта, поверхонь або країв сцени, що отримується в результаті переміщення спостерігача (очей або камери) щодо сцени. Алгоритми, засновані на оптичному потоці (реєстрація руху, сегментація об'єктів, кодування рухів) використовують цей рух об'єктів, поверхонь і країв. По суті, оптичний потік - це поле швидкостей, так як зсув точки між двома зображеннями еквівалентний миттєвої швидкості. Обчислення оптичного потоку - знаходження векторів (швидкостей) для оцінки руху об'єкта [26].

Застосування даного підходу вимагає виконання двох основних умов:

- яскравість кожної точки об'єкта не змінюється з плином часу;
- найближчі точки, що належать одному об'єкту, в площині зображення рухаються зі схожою швидкістю.

2.3.1. Оцінка оптичного потоку

Методи, засновані на оптичному потоці, обчислюють рух між двома кадрами в кожному пікселі. Кадри беруться в момент часу t та $t + \delta t$. Дані методи називаються диференціальними, так як вони засновані на наближенні сигналу відрізком ряду Тейлора. Тому ці підходи застосовують приватні похідні по часу і просторовим координатам.

У разі розмірності $2D + t$ (при більшій розмірності аналогічно) піксель в позиції (x, y, t) з інтенсивністю $I(x, y, t)$ за один кадр буде переміщений на δx , δy і δt . Внаслідок того, що інтенсивність точки не змінюється, виходить:

$$I(x, y, t) = I(x + \delta x, y + \delta y, t + \delta t)$$

Так як переміщення мале, за допомогою ряд Тейлора виводиться рівняння:

$$I(x + \delta x, y + \delta y, t + \delta t) \approx I(x, y, t) + \frac{\partial I}{\partial x} \delta x + \frac{\partial I}{\partial y} \delta y + \frac{\partial I}{\partial t} \delta t$$

З даних рівнянь виходить:

$$\frac{\partial I}{\partial x} \delta x + \frac{\partial I}{\partial y} \delta y + \frac{\partial I}{\partial t} \delta t = 0$$

або

$$\frac{\partial I}{\partial x} \frac{\delta x}{\delta t} + \frac{\partial I}{\partial y} \frac{\delta y}{\delta t} + \frac{\partial I}{\partial t} \frac{\delta t}{\delta t} = 0$$

Виходячи з цього отримуємо:

$$\frac{\partial I}{\partial x} V_x + \frac{\partial I}{\partial y} V_y + \frac{\partial I}{\partial t} = 0,$$

де v_x, v_y – компоненти швидкості оптичного потоку в $I(x, y, t)$,

$\frac{\partial I}{\partial x} \frac{\partial I}{\partial y} \frac{\partial I}{\partial t}$ – похідні зображення в (x, y, t) у відповідних напрямках.

Таким чином:

$$I_x V_x + I_y V_y = -I_t$$

або

$$\nabla I^T \cdot \bar{V} = -I_t$$

Це рівняння оптичного потоку [27]. Дане рівність містить дві невідомі V_x, V_y і не може бути однозначно вирішеною. Отже, необхідно ввести додаткове припущення. Наприклад, нехай оптичний потік змінюється плавно від кадру до кадру, тобто для всіх пікселів p з околиці (x, y) оптичного потоку зміщення буде постійним.

2.3.2. Методи визначення оптичного потоку

Існує безліч різних методів визначення оптичного потоку:

- Фазова кореляція - інверсія нормалізованого перехресного спектра.
- Блокові методи - пошук місця розташування заданих областей (Блоків) поточного кадру на попередньому кадрі.
- Диференціальні методи оцінки оптичного потоку,

базуються на приватних похідних сигналу:

1. Алгоритм Лукаса - Канаді
2. Horn-Schunck
3. Buxton-Buxton

Horn-Schunck.

При використанні алгоритму Horn-Schunck враховується, що на всьому зображенні оптичний потік буде досить гладким [28]. Вводяться дві функції, що визначають помилку:

1. Характеризує розбіжність інтенсивностей вихідного і зміщеного пікселів

$$E_d(x, u) = E_d(I_0(x) - I_1(x + u)) = |I_0(x) - I_1(x + u)|^2 = \left| u \cdot \nabla I(x, t) + \frac{\partial I(x, t)}{\partial t} \right|^2,$$

де u - зміщення точки.

2. Характеризується ще одна вимога на відсутність різкого зміни швидкостей/зсувів.

$$E_s(u) = E_s(u, \nabla u) = |\nabla u|^2$$

Відбувається мінімізація аналога енергії за формулою:

$$E = \int_{x \in I} E_d + \lambda E_s dx = \int_{x \in I} \left| u \cdot \nabla I(x, t) + \frac{\partial I(x, t)}{\partial t} \right|^2 + \lambda |\nabla u(x)|^2 dx,$$

де λ - невід'ємні ваговий коефіцієнт.

В даному випадку може бути застосовано рівняння Ейлера-Лагранжа:

$$I_x^2 u + I_x I_y v = \lambda \nabla^2 u - I_x I_t$$

$$I_x I_y u + I_y^2 v = \lambda \nabla^2 v - I_y I_t$$

Виходячи з даного рівняння, наблизивши лапласіан зваженою сумою і виписавши такі рівняння для всіх точок зображення, можна знайти $u(x)$.

TV-L1 (Total Variation in L1).

Підходи, що базуються на методі найменших квадратів, працюють ефективно, коли помилка в збереженні інтенсивності точки:

$$e(x) = u \nabla I(x, t) + I_t(x, t)$$

розподілена приблизно як $N(0, \sigma^2)$ [26]. Дане твердження виконується рідко, так як інтенсивність об'єктів змінюється. Це відбувається через тіней, що залежать від часу доби, мерехтінь освітлення, змін нахилу поверхонь, появи відблисків і так далі. Дані методи неефективні в тих випадках, коли $e(x)$ добре описується розподілами з так званими «важкими хвостами», також вони не стійкі до викидів [26].

В основі алгоритму DTV-L1 лежить та ж ідея, що і в алгоритмі HornSchunk, але має в основі підінтегральної функції в метриці L1, тобто

$$E_d(u \cdot \nabla I(x, t) + \frac{\partial I(x, t)}{\partial t}) = \left| u \cdot \nabla I(x, t) + \frac{\partial I(x, t)}{\partial t} \right|$$

$$E_s(u, \nabla u) = |\nabla u|$$

У разі використання функцій такого виду виникають деякі математичні складності: чи не диференційність модуля в 0, тобто неможливість використання формули Ейлера-Лагранжа для варіації функціонала $\frac{\delta E}{\delta x}$. Для вирішення даної проблеми було вирішено модифікувати підінтегральної функції, приводячи її до вигляду, зручного для оптимізації [30].

MDPOF (Motion Detail Preserving Optical Flow Estimation).

Існує алгоритм, який застосовує мінімізацію функціоналу "енергії" в метриці L1, який вводить додаткове бінарне поле $\gamma(x) \in \{0, 1\}$, $x \in \Pi$, що визначає заслонення великого об'єкта малим. [31]

$$E_d(x, u, \gamma) = \gamma |I_1(x+u) - I_0(x)| + (1-\gamma) |\nabla I_1(x+u) - \nabla I_0(x)|$$

$$E_s(x, u) = \omega(x) |\nabla u|,$$

де $\omega(x) = \exp(-|\nabla I_0(x)|^k)$, $k \sim 0.8$

$$E(u) = \int_{x \in \Omega} E_d(x, u, \gamma(x)) + \lambda E_s(x, u) dx$$

Підхід, який обчислюється бінарне поле $\gamma(x)$, вирішує проблему коректної обробки країв об'єкта, а саме місць, де $\Delta u(x)$ терпить розрив [31].

Метод Лукаса–Канаде.

Алгоритм Лукаса-Канаде є одним з методів, що застосовуються для обчислення оптичного потоку. Даний алгоритм використовується досить широко в задачах оцінки руху об'єкта. Метод Лукаса-Канаде обробляє пікселі в околиці конкретної точки, тому його називають локальним підходом обчислення оптичного потоку [32].

Алгоритм передбачає, що:

1. Зміщення точок на двох кадрах відеопослідовності незначне.
2. Зміщення точок в околиці деякої точки p однаково.

Існує припущення: $D = \{q_1, q_2, \dots, q_n\}$ - набір точок в околиці точки p .

Тоді, можна записати систему рівнянь:

$$\begin{cases} I_x(q_1)V_x + I_y(q_1)V_y = -I_t(q_1) \\ I_x(q_2)V_x + I_y(q_2)V_y = -I_t(q_2) \\ \dots \\ I_x(q_n)V_x + I_y(q_n)V_y = -I_t(q_n) \end{cases},$$

де q_1, q_2, \dots, q_n – точки всередині вікна пошуку,

$I_x(q_i), I_y(q_i), I_t(q_i)$ – приватні похідні зображення I по x, y і часу t ,

що обчислюються в точці q_i в поточний момент часу.

Ці вирази можна переписати в матричній формі $Av = b$, де

$$A = \begin{bmatrix} I_x(q_1) & I_y(q_1) \\ I_x(q_2) & I_y(q_2) \\ \vdots & \vdots \\ I_x(q_n) & I_y(q_n) \end{bmatrix}, \quad v = \begin{bmatrix} V_x \\ V_y \end{bmatrix}, \quad b = \begin{bmatrix} -I_t(q_1) \\ -I_t(q_2) \\ \vdots \\ -I_t(q_n) \end{bmatrix}$$

Ця система рівнянь зазвичай сильно надлишкова, тому зазвичай її вирішують методом найменших квадратів. В результаті виходить:

$$A^T Av = A^T b$$

Виводимо:

$$v = (A^T A)^{-1} A^T b,$$

де A^T – транспонована матриця A .

Цей вираз можна записати в матричному вигляді:

$$\begin{bmatrix} V_x \\ V_y \end{bmatrix} = \begin{bmatrix} \sum_i I_x(q_i)^2 & \sum_i I_x(q_i)I_y(q_i) \\ \sum_i I_x(q_i)I_y(q_i) & \sum_i I_y(q_i)^2 \end{bmatrix}^{-1} \begin{bmatrix} -\sum_i I_x(q_i)I_t(q_i) \\ -\sum_i I_y(q_i)I_t(q_i) \end{bmatrix}$$

В основному набір точок в околиці точки p обмежений вікном, розмір якого $N \times M$. Основна проблема полягає в прямій залежності між кількістю рівнянь в системі і числом точок в околиці певної точки. Тобто, чим більше кількості точок в околиці точки, тим більше рівнянь буде містити система.

Для вирішення проблеми використовується зважений метод найменших квадратів. При визначенні вагових коефіцієнтів пікселів на зображенні застосуються функція $W(x,y)$ [33]. Виходячи з даного методу, для знаходження рішення необхідно мінімізувати нев'язку:

$$\begin{aligned} E(V) &= \sum_{x,y \in D} W(x,y) [I(x,y,t) - I(x+\delta x, y+\delta y, t+\delta t)]^2 = \\ &= \sum_{x,y \in D} W(x,y) \left(\frac{\partial I}{\partial x} V_x + \frac{\partial I}{\partial y} V_y + \frac{\partial I}{\partial t} \right)^2 \end{aligned}$$

де $V = (V_x, V_y)$

$W(x,y)$ - вагові коефіцієнти, присвоєні пікселям q_i .

Щоб знайти мінімум $E(V)$ застосовують метод найменших квадратів, з перебуванням її приватних похідних по V_x та V_y :

$$\begin{aligned} \frac{\partial E(V)}{\partial V_x} &= 2 \sum_{x,y \in D} W(x,y) \left(\frac{\partial I}{\partial x} V_x + \frac{\partial I}{\partial y} V_y + \frac{\partial I}{\partial t} \right) \frac{\partial I}{\partial x} = \\ &= 2 \sum_{x,y \in D} W(x,y) \left[\left(\frac{\partial I}{\partial x} \right)^2 V_x + \frac{\partial I}{\partial y} \frac{\partial I}{\partial x} V_y + \frac{\partial I}{\partial t} \frac{\partial I}{\partial x} \right] \end{aligned}$$

$$\begin{aligned} \frac{\partial E(V)}{\partial V_y} &= 2 \sum_{x,y \in D} W(x,y) \left(\frac{\partial I}{\partial x} V_x + \frac{\partial I}{\partial y} V_y + \frac{\partial I}{\partial t} \right) \frac{\partial I}{\partial y} = \\ &= 2 \sum_{x,y \in D} W(x,y) \left[\frac{\partial I}{\partial x} \frac{\partial I}{\partial y} V_x + \left(\frac{\partial I}{\partial y} \right)^2 V_y + \frac{\partial I}{\partial t} \frac{\partial I}{\partial y} \right] \end{aligned}$$

Дані рівняння можна переписати в більш компактній формі і прирівняємо до нуля:

$$\begin{cases} \sum_{x,y \in D} W(x,y) \left[\left(\frac{\partial I}{\partial x} \right)^2 V_x + \frac{\partial I}{\partial y} \frac{\partial I}{\partial x} V_y + \frac{\partial I}{\partial t} \frac{\partial I}{\partial x} \right] = 0 \\ \sum_{x,y \in D} W(x,y) \left[\frac{\partial I}{\partial x} \frac{\partial I}{\partial y} V_x + \left(\frac{\partial I}{\partial y} \right)^2 V_y + \frac{\partial I}{\partial t} \frac{\partial I}{\partial y} \right] = 0 \end{cases}$$

Система рівнянь може бути представлена в матричній формі:

$$\begin{bmatrix} \sum_{x,y \in D} W(x,y) \left(\frac{\partial I}{\partial x} \right)^2 & \sum_{x,y \in D} W(x,y) \left(\frac{\partial I}{\partial y} \frac{\partial I}{\partial x} \right) \\ \sum_{x,y \in D} W(x,y) \left(\frac{\partial I}{\partial x} \frac{\partial I}{\partial y} \right) & \sum_{x,y \in D} W(x,y) \left(\frac{\partial I}{\partial y} \right)^2 \end{bmatrix} \begin{bmatrix} V_x \\ V_y \end{bmatrix} + \begin{bmatrix} \sum_{x,y \in D} W(x,y) \left(\frac{\partial I}{\partial t} \frac{\partial I}{\partial x} \right) \\ \sum_{x,y \in D} W(x,y) \left(\frac{\partial I}{\partial t} \frac{\partial I}{\partial y} \right) \end{bmatrix} = 0$$

Нехай

$$\begin{aligned} A &= \begin{bmatrix} \sum_{x,y \in D} W(x,y) \left(\frac{\partial I}{\partial x} \right)^2 & \sum_{x,y \in D} W(x,y) \left(\frac{\partial I}{\partial y} \frac{\partial I}{\partial x} \right) \\ \sum_{x,y \in D} W(x,y) \left(\frac{\partial I}{\partial x} \frac{\partial I}{\partial y} \right) & \sum_{x,y \in D} W(x,y) \left(\frac{\partial I}{\partial y} \right)^2 \end{bmatrix} \quad V = \begin{bmatrix} V_x \\ V_y \end{bmatrix} \\ B &= \begin{bmatrix} \sum_{x,y \in D} W(x,y) \left(\frac{\partial I}{\partial t} \frac{\partial I}{\partial x} \right) \\ \sum_{x,y \in D} W(x,y) \left(\frac{\partial I}{\partial t} \frac{\partial I}{\partial y} \right) \end{bmatrix} \end{aligned}$$

Тоді

$$A \cdot V + B = 0$$

або

$$V = A^{-1} B$$

Цей вираз можна записати як систему рівнянь:

$$\begin{bmatrix} V_x \\ V_y \end{bmatrix} = \begin{bmatrix} \sum_{x,y \in D} W(x,y) \left(\frac{\partial I}{\partial x} \right)^2 & \sum_{x,y \in D} W(x,y) \left(\frac{\partial I}{\partial y} \frac{\partial I}{\partial x} \right) \\ \sum_{x,y \in D} W(x,y) \left(\frac{\partial I}{\partial x} \frac{\partial I}{\partial y} \right) & \sum_{x,y \in D} W(x,y) \left(\frac{\partial I}{\partial y} \right)^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{x,y \in D} W(x,y) \left(\frac{\partial I}{\partial t} \frac{\partial I}{\partial x} \right) \\ \sum_{x,y \in D} W(x,y) \left(\frac{\partial I}{\partial t} \frac{\partial I}{\partial y} \right) \end{bmatrix}$$

В даний час методи обчислення оптичного потоку стають дедалі популярнішими. Їх застосовують при визначенні напрямку руху об'єктів; визначенні відстані між об'єктами з допомогою аналізу оптичного потоку кадрів, отриманих з двох камер (стереозрення). За допомогою методів оптичного потоку може бути проведено найбільш акуратне виділення рухомих об'єктів, навіть в разі переміщення камери.

Висновки за розділом 2

Підсумовуючи другий розділ, можемо зробити такі висновки:

1. Розглянуто основні групи методів розпізнавання рухів об'єкту.
2. Визначено, що у відповідності до мети дослідження, доцільно використовувати структурні способи розпізнавання.
3. Окреслено існуючі інструменти розпізнавання образів, які використовуються, а також з'ясовані їх переваги та недоліки.

3 РОЗРОБКА ТА ДОСЛІДЖЕННЯ АЛГОРИТМІВ РОЗПІЗНАВАННЯ РУХОМИХ ОБ'ЄКТІВ

Для розробки і дослідження були обрані наступні методи:

- метод вирахування фону;
- метод вирахування фону з використанням суміші Гауссових розподілів (GMM);
- метод оптичного потоку (алгоритм Лукаса-Канаде).

Реалізація та дослідження обраних методів було проведено в програмному комплексі MATLAB і Simulink з використанням пакета Computer Vision System Toolbox.

Computer Vision System Toolbox (набір засобів для проектування систем комп'ютерного зору) містить алгоритми та інструменти для розробки і моделювання систем комп'ютерного зору і обробки відео. Він включає методи для виявлення деталей, визначення руху, виявлення та відстеження об'єктів, стереозору, обробки і аналізу відео. Є інструменти для введення виведення відео-файлів, відображення відео, побудови графіки та компоновання.

Дані можливості надані у вигляді системних об'єктів і функцій MATLAB і блоків Simulink. Крім того, системний набір інструментів підтримує арифметику з фіксованою точкою і генерацію Сі коду, що дозволяє швидко створювати прототипи системи і розробляти вмонтовані системи.

Були розроблені алгоритми обраних методів:

- метод вирахування фону;
- метод вирахування фону з використанням суміші Гауссових розподілів (GMM);
- алгоритм Лукаса Канаді для обчислення оптичного потоку.

3.1. Розробка алгоритмів виявлення рухомих об'єктів

3.1.1. Розробка алгоритму віднімання фону

Був розроблений алгоритм віднімання фону. Блок-схема алгоритму представлена на малюнку 3.1.

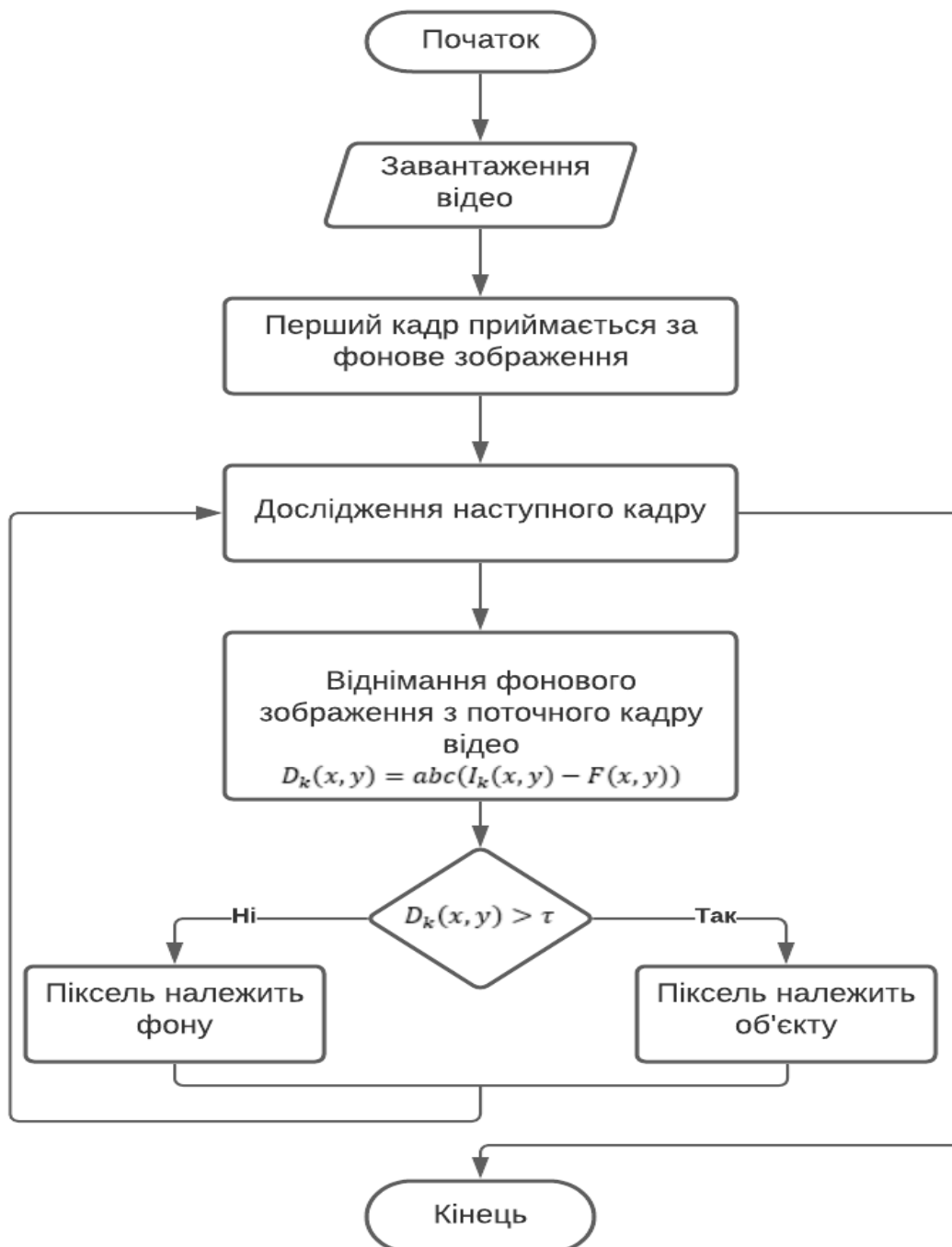


Рисунок 3.1 - Блок-схема алгоритму віднімання фону

Блок 1 - Початок;

Блок 2 - Введення відео-даних;

Блок 3 - Перший кадр відео приймається за фонове зображення без рухомих об'єктів;

Блок 4 - Дослідження наступного кадру;

Блок 5 - Знаходиться різниця поточного кадру відео і фонового зображення;

Блок 6 - Визначення істинності виразу (чи більша отримана різниця обраного порога): $D_k(x, y) > \tau$

- якщо істинно, тобто $(D_k(x, y) > \tau)$, то здійснюється перехід до блоку 6;
- якщо помилково, тобто $(D_k(x, y) < \tau)$, то здійснюється перехід до блоку 7;

Блок 7 - Піксель визначається, як той, що належить об'єкту;

Блок 8 - Піксель визначається, як той, що належить фону;

Блок 9 - Кінець.

3.1.2. Розробка алгоритму віднімання фону з використанням суміші Гауссових розподілів (GMM)

Був розроблений алгоритм віднімання фону з використанням суміші Гауссових розподілів (GMM). Блок-схема алгоритму представлена на малюнку 3.2.

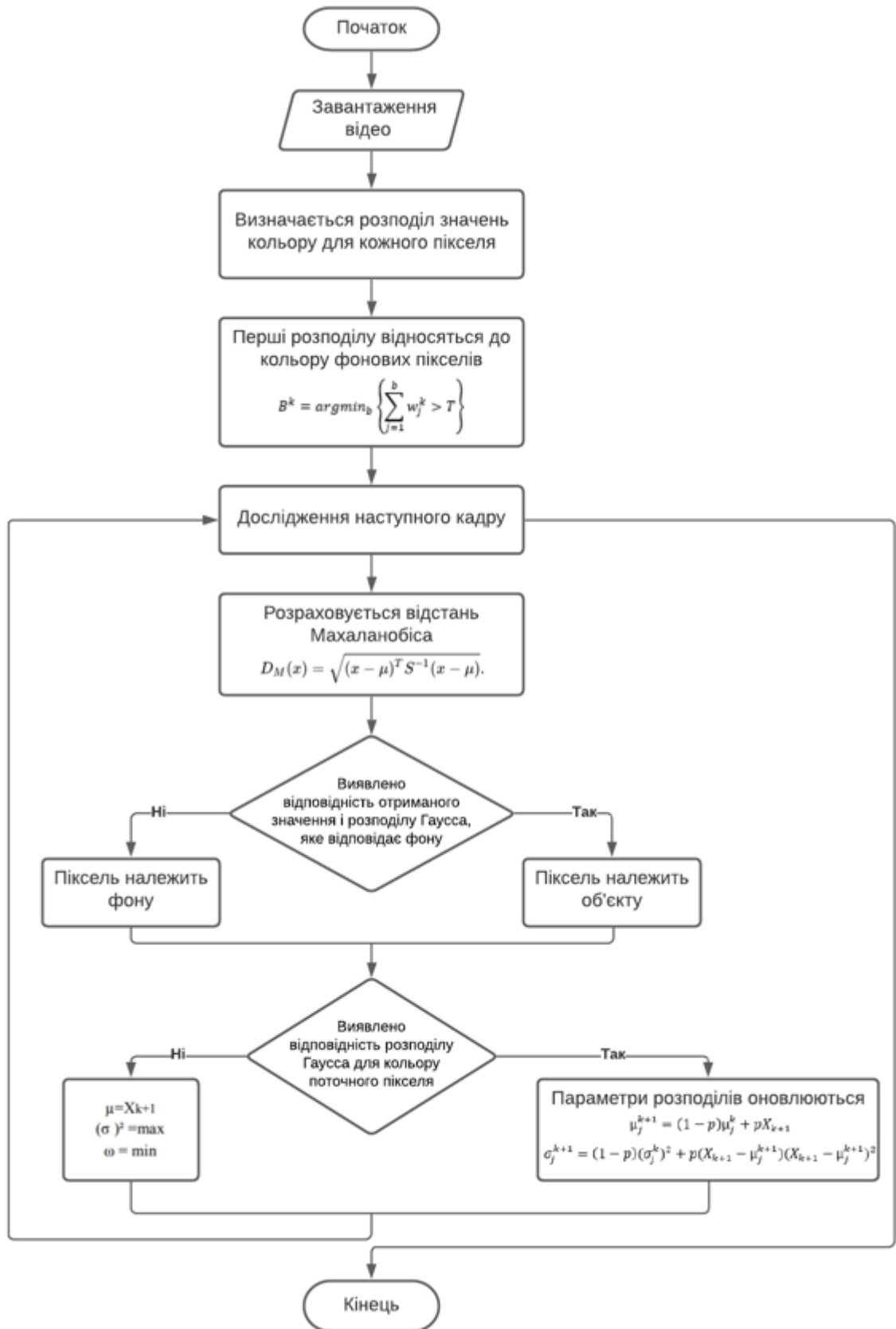


Рисунок 3.2 - Блок-схема алгоритму віднімання фону, заснованого на суміші Гауссових розподілів

Блок 1 - Початок;

Блок 2 - Введення відео-даних;

Блок 3 - Визначається розподіл спостережуваних значень кольору для кожного пікселя. Нове значення кольору можна уявити як одну з основних компонент отриманої суміші Гауссових розподілів.

Блок 4 - Перші B^k розподілів відносяться до кольору фонових пікселів, коли виконується така умова:

$$B^k = \operatorname{argmin}_b \{ \sum_{j=1}^b w_j^k > T \},$$

де T - порогове значення, параметр моделі.

Блок 5 - Дослідження наступного кадру;

Блок 6 - Для кожного пікселя зображення виконується тест, за допомогою якого при використанні відстані Махаланобіса, визначається, яким розподілу відповідає отримане значення.

Блок 7 - Перевіряється виявлено чи відповідність отриманого значення і розподілу Гаусса, відповідного фону:

- якщо відповідність виявлено, то піксель визначається як фоновий;

- якщо відповідність не виявлено, то приймається, що піксель належить об'єкту.

Блок 8 - Перевіряється виявлено чи відповідність розподілу Гаусса для кольору поточного пікселя:

- якщо відповідність виявлено, то параметри розподілів оновлюються за такими формулами:

$$\mu_j^{k+1} = (1 - p)\mu_j^k + pX_{k+1}$$

$$\sigma_j^{k+1} = (1 - p)(\sigma_j^k)^2 + p(X_{k+1} - \mu_j^{k+1})(X_{k+1} - \mu_j^{k+1})^2$$

де $p = \alpha N(X_{k+1} | \mu_j^k, \Sigma_j^b)$.

- якщо відповідність не виявлено, то математичне очікування зрівнюється поточного значення кольору пікселя. Дисперсія $(\sigma_s^{k+1})^2$ обирається максимально можливою, а вага w_s^{k+1} - мінімально допустимою.

Блок 7 - Кінець.

3.1.3. Розробка алгоритму Лукаса-Канаде для обчислення оптичного потоку

Був розроблений алгоритм Лукаса-Канаде для обчислення оптичного потоку. Блок-схема алгоритму представлена на малюнку 3.3.

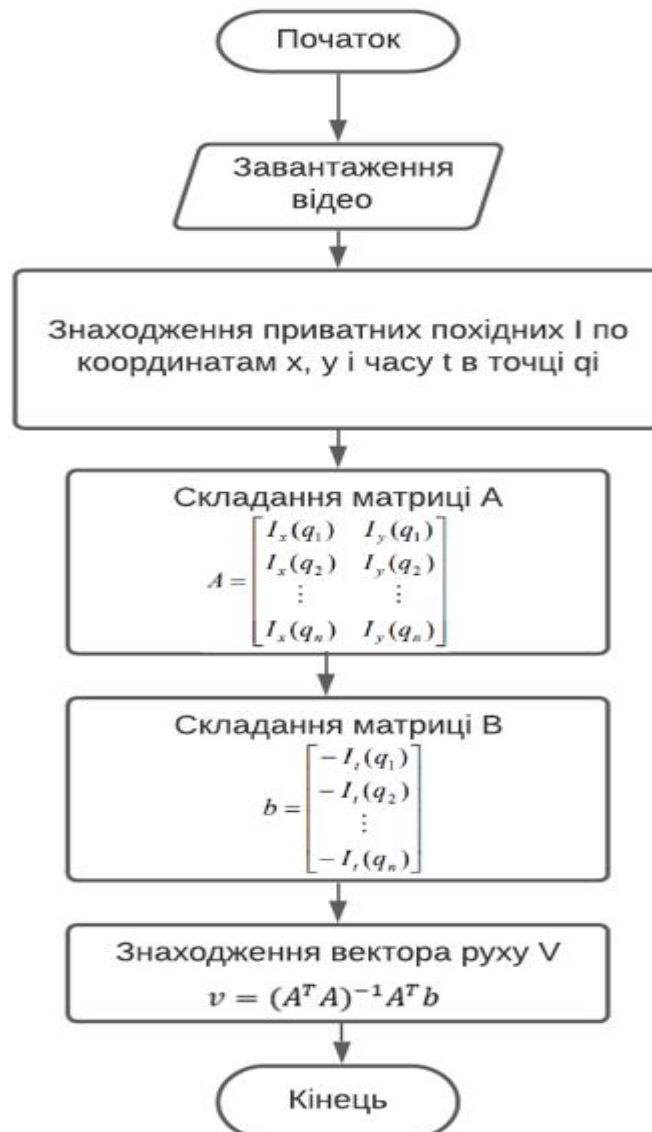


Рисунок 3.3 - Блок-схема алгоритму Лукаса-Канаді для обчислення оптичного потоку

Блок 1 - Початок;

Блок 2 - Введення відео-даних;

Блок 3 - Знаходяться приватні похідні зображення I по x , y і часу t , що обчислюються в точці q_i в поточний момент часу;

Блок 4 - Складається матриця A , що складається з приватних похідних зображення I по x , y , обчислювані в точці q_i ;

Блок 5 - Складається матриця b , що складається з приватних похідних зображення I за часом t , обчислювані в точці q_i ;

Блок 6 - Знаходиться вектор руху;

Блок 7 - Кінець.

3.2. Дослідження алгоритмів виявлення рухомих об'єктів

Для дослідження алгоритмів були обрані 4 відео-фрагмента. При виборі відеороликів враховувалися наступні вимоги: відео має бути знято в різних навколишніх умовах, в різний час доби, повинні бути різні погодні умови.

Відповідно до даних вимог були використані відео, опис яких представлено в таблиці 3.1

Таблиця 3.1 - Опис використаних відеороликів

Назва відео	Тип відео	Що в кадрі	Умови зйомки
1	2	3	4
day_cars.avi	720x576, 25fps	Машини	Сонячний день
artificial_light.avi	720x576, 25fps	Люди	День, штучне освітлення приміщення
night_cars.avi	720x576, 25fps	Дорога, машини	Ніч, світло від фар автомобілів
blizzard.avi	720x576, 25fps	Машини, люди	Похмурий день, сніг

Було проведено дослідження розроблених алгоритмів. Результати досліджень представлені на малюнках 3.4 - 3.7.



а)



б)



в)

Рисунок 3.4 - Результати роботи алгоритмів:

а) Алгоритм віднімання фону

б) Алгоритм Лукаса-Канаді

в) Алгоритм віднімання фону з використанням GMM

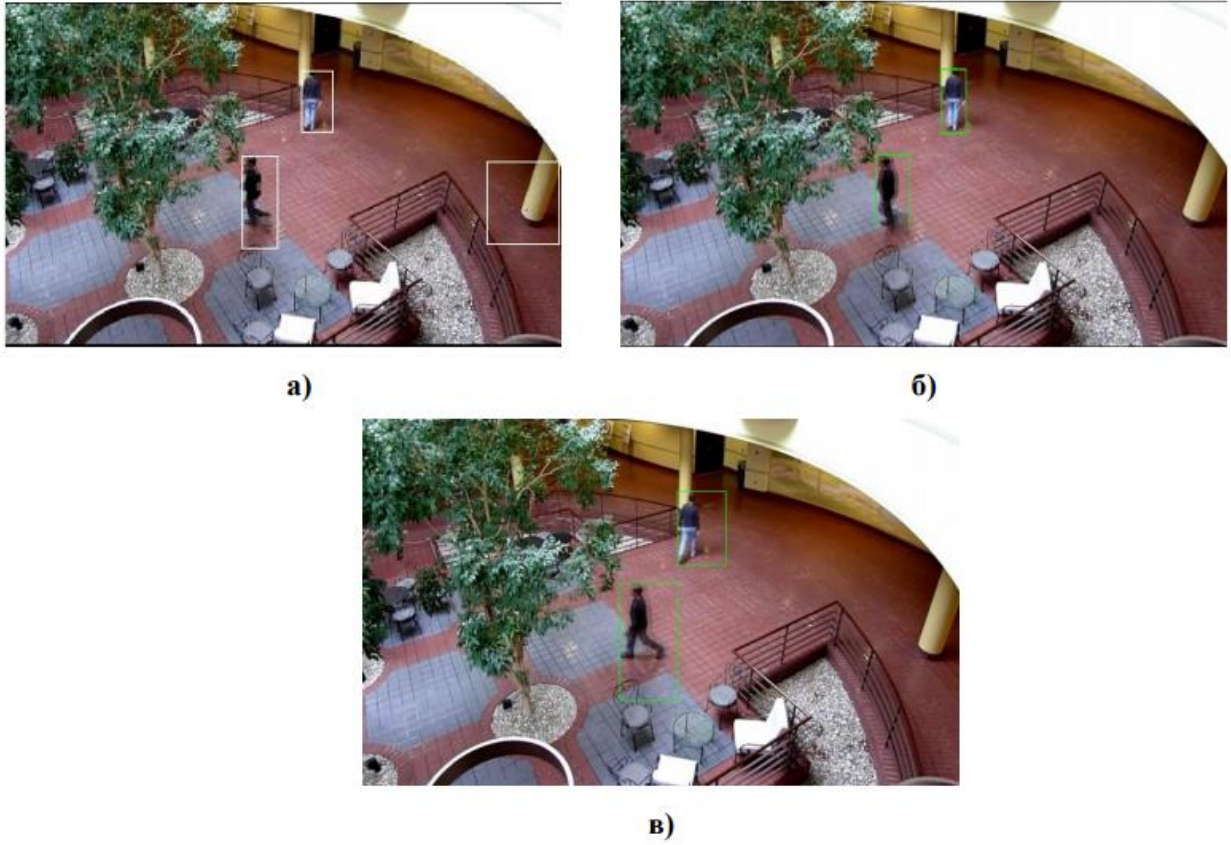


Рисунок 3.5 - Результати роботи алгоритмів:

а) Алгоритм віднімання фону

б) Алгоритм Лукаса-Канаде

в) Алгоритм віднімання фону з використанням GMM



а)

б)



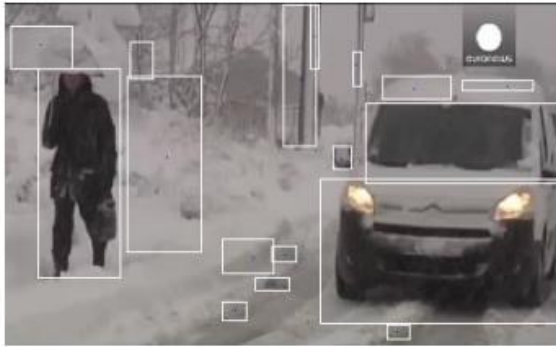
в)

Рисунок 3.6 - Результати роботи алгоритмів:

а) Алгоритм віднімання фону

б) Алгоритм Лукаса-Канаде

в) Алгоритм віднімання фону з використанням GMM



а)



б)



в)

Рисунок 3.7 - Результати роботи алгоритмів:

а) Алгоритм віднімання фону

б) Алгоритм Лукаса-Канаде

в) Алгоритм віднімання фону з використанням GMM

За малюнком 3.4 можна сказати, що алгоритм віднімання фону (рисунок 3.4 (а)) не визначив рух другої машини, яка тільки почала свій рух. Це пояснюється латентністю в оновленні моделі фону: має пройти деякий час, перш ніж в моделі будуть враховані зміни, пов'язані з початком руху або зупинкою об'єкта. Це можна спостерігати і у алгоритму віднімання фону з використанням суміші Гауссових розподілів (рисунок 3.4 (в)). Однак метод Лукаса-Канаде виявляє рухомі об'єкти без тимчасових затримок (рисунок 3.4 (б)).

Алгоритм віднімання фону і алгоритм з використанням суміші Гауссових розподілів не можуть відокремити рухомі тіні від об'єктів. Тому на малюнках 3.5 (а) і 3.5 (в) разом з об'єктом виділяється і область з тінню. Тоді як алгоритм Лукаса-Канаде не захоплює тінь при виділенні об'єкта, тобто виділення більш точне (рисунок 3.5 (б)).

Алгоритм віднімання фону і алгоритм з використанням суміші Гауссових розподілів досить неефективні при мінімальному освітленні. Виникають помилкові спрацьовування: виділяються світлі області машин і освітлені ділянки дороги. Рух самих об'єктів не визначається (малюнки 3.6 (а, в)). Алгоритм Лукаса-Канаде не залежить від рівня освітленості. Що показано на малюнку 3.6 (б).

Несприятливі погодні умови впливають на роботу алгоритмів віднімання фону. За малюнками 3.7 (а) і 3.7 (в) можна сказати, що падаючий сніг викликає помилкові спрацьовування. При роботі алгоритму віднімання фону з використанням суміші Гауссових розподілів спостерігається менша кількість помилкових спрацьовувань, ніж при роботі алгоритму віднімання фону. Алгоритм Лукаса-Канаде стійкий до перешкод такого роду. Однак даний метод має сильні обмеження в допустимій швидкості руху об'єктів - рух повільніший, ніж один піксель за кадр, не виявляється зовсім (рисунок 3.7 (б)).

Висновки за розділом 3

Підсумовуючи третій розділ, можна зробити наступні висновки:

1. Представлено структурні схеми алгоритмів обробки і аналізу зображень та проаналізовані функціональні можливості їх компонентів;
2. Розроблені алгоритми Лукаса-Канаде для обчислення оптичного потоку, алгоритм віднімання фону, а також алгоритм віднімання фону з використанням суміші Гауссових розподілів (GMM).;
3. Досліджена робота алгоритмів виявлення рухомих об'єктів за допомогою програмного середовища Matlab та чотирьох відеороликів, які зняті при різних навколишніх умовах, в різний час доби та при різних погодних умовах.

ВИСНОВКИ

Під час виконання бакалаврської роботи було проведено огляд кількох методів виявлення рухомих об'єктів на відеопослідовності. Можна відзначити, що існує безліч методів виявлення об'єктів: алгоритми, що базуються на моделюванні та відніманні заднього фону, а також тимчасових відмінностях кадрів тощо.

Відповідно, для дослідження було обрано наступні методи виявлення рухомих об'єктів на відео-даних:

- метод вирахування фону з використанням суміші Гауссова розподілів;
- метод вирахування фону;
- метод оптичного потоку (алгоритм Лукаса-Канаде).

Були розроблені алгоритми цих методів, виконана програмна реалізація в програмному пакеті Matlab, після цього були проведені дослідження розроблених алгоритмів при різних умовах освітленості та погодних умовах.

Виходячи з отриманих результатів дослідження можна зробити наступні висновки:

- Алгоритм Лукаса-Канаде для побудови оптичного потоку досить стійкий до змін освітленості, яскравості і контрастності кадрів відеопослідовності. Недоліками даного методу є нестійкість до впливу шуму, а також неефективність виявлення об'єктів при повільних рухах (повільніше, ніж 1 піксель за кадр). Даний алгоритм також є більш стійким до змін розширення відео-даних, ніж алгоритми віднімання фону.
- Головним недоліком методу віднімання фону є можлива класифікація фонових пікселів як передньопланових. Це виникає при появі на відеопослідовності листя дерев, що колишуться від вітру, падаючого снігу і дощу, відбитих рухомими об'єктами

тінями, тощо. Для даного методу також властива латентність в оновленні моделі фону. Алгоритм не стійкий до зміни розширення відео-даних, тобто при зменшенні розширення відео - ймовірність правильного виявлення об'єктів, що рухаються значно падає. Переваги даного методу: достатня стійкість до впливу шуму будь-якої інтенсивності та умовна простота реалізації.

ПЕРЕЛІК ПОСИЛАНЬ

1. Гренандер У. Лекції по теорії образів. Т.1. Синтез образів. – М.: Мир, 1979. – 382с.
2. Антонюк В.С., Вислоух С.П., Аверченков В. І. - Автоматизоване проектування технологічних процесів. – Київ, УМК ВО, 1989 – 116 с.
3. Главач В., Шлезингер М. Десять лекцій по статистичному і структурному розпізнанню образів. К.: Наукова думка, 2004. -Режим доступу: www.irtc.org.ua/image/Files/Schles/esh10_full.pdf.
4. Вислоух С. П. Інформаційні технології в задачах технологічної підготовки приладо- та машинобудівного виробництва: монографія/С. П. Вислоух. – К. : НТУУ «КПІ», 2011. – 488 с.
5. Дейч А. М. Методи ідентифікації динамічних об'єктів. – М: Енергія, 1979 – 240 с.
6. Антонюк В.С., Вислоух С.П., Катрук О.В. Класифікації розпізнання образів при автоматизованому проектуванні технологічних процесів. // Надійність інструменту та оптимізація технологічних систем. Збірник наукових праць. – Краматорськ–Київ, Вип. №23, 2008. – С. 176–182.
7. Терехов В. А., Єфімов Д. В., Тюкин І. Ю. Нейромережні системи керування. - 1-е. - Вища школа, 2002. - С. 184.
8. Глушков В.М., Амосов М.М., Артеменко І. О. Енциклопедія кібернетики. Том 2. Київ, 1974. – С. 12-20.
9. Дремін І. М., Иванов О. В., Нечитайло В. О. Вейвлетитаїх використання //Успіхи фізичних наук. - 2001. - Т. 171. - № . 5. - С. 465-501.
10. Аваліані Г. В. Евристичні методив розпізнаванні образів / Г. В.Аваліані. – Тбілісі: Мецніереба, 1988. – 78 с.
11. Бібліотека OpenCV [Електронний ресурс]. - Режим доступу: <https://uk.wikipedia.org/wiki/OpenCV>.

12. Вудс Р., Гонсалес Р. Цифрове оброблення зображень //М.: Техносфера. - 2005. – С. 44-58.
13. Горелик О. Л. Методи розпізнання / О. Л. Горелик, В. О. Скрипкин. – М.: ВШ, 1984, 2004. – 262 с.
14. Основи теорії розпізнавання образів : навч. посіб. : у 2 ч. / А. С. Довбиш, І. В. Шелехов. – Суми : Сумський державний університет, 2015. – Ч. 1. – 109 с.
15. Патрік Е. Основи теорії розпізнання образів. - М.: Радянське Радіо, 1980. - 408 с.
16. Роберт Каллан. Основні концепції нейронних мереж = The Essence of Neural Networks First Edition. - 1-е. - «Вільямс», 2001. - С. 288.
17. Ясницький Л. М. Введення в штучний інтелект. - 1-е видання. – Видавничий центр «Академія», 2005. - С. 176.
18. Шапіро Л., Стокман Д. Комп'ютерний зір//М.: Біном. Лабораторія знань. - 2006. - Т. 752.
19. Уосермен Ф. Нейрокомп'ютерна техніка: Теорія і практика. Переклад українською І.Ю.Юрчак, 2001 р. – С. 56-82.
20. Рутковська Д. Нейронні мережі, генетичні алгоритми та нечіткі системи / Д. Рутковська, М. Пилинський, Л. Рутковський. - М. Гаряча лінія-Телеком 2004. – 112 с.
21. Советов Б.Я. Інформаційні технології / Б.Я. Советов, В.В. Цехановський - М.: Вища школа, 2005 - 263 с.
22. Тархов Д.О. Нейронні мережі. Моделі та алгоритми. – М.: Радіотехніка, 2010. – 82 с.
23. Ту Дж., Гонсалес Р. Принципи розпізнавання образів. – М.: Мир, 1978 – 412с.
24. Фомін Я. О. Розпізнання образів: теорія та застосування / Я. О. Фомін. – М.: ФАЗИС, 2012. – 429 с.
25. Tarasov A., Potapova V., Belyakov P., Melnik O. The Developing of Targets Tracking Complex // Proceedings of EWDTTS-2018, Казань, 2018.

26. Беляков П.В., Никифоров М.Б. Варіаційний метод обчислення оптичного потоку в системі-на-кристалі // Цифрова обробка сигналів, 2018. №3. С. 76-82.

27. Новиков А.І., Сабліна В.А., Горячев Е.О. Застосування контурного аналізу для суміщення зображень // Звістки Тульського державного університету. Технічні науки, 2013. Вип. 9. Ч. 1. С. 260 - 270.

28. Абдухаліков А., Беляков П., Никифоров М. Реалізація на ПЛІС алгоритму пошуку ключових точок на зображенні // МНТК «Сучасні технології в науці та освіті». Рязань, 2016. С. 103-108.

29. Tarasov A., Potapova V., Belyakov P., Melnik O. The Developing of Targets Tracking Complex // Proceedings of EWDTS-2018, Казань, 2018.

30. Lucas B.D., Kanade T. An iterative image registration technique with an application to stereo vision // Proceedings of Imaging Understanding Workshop, 1981. С. 121-130.

31. Ustukov D., Muratov Y., Nikiforov M., Gurov V. Implementing one of stereovision algorithms on FPGA // Mediterranean Conference on Embedded Computing. Bar, 2016. P. 125-130.