ДИПЛОМНА РОБОТА
БАКАЛАВРА

ДРБ. ІП - 46.00.00.000 ПЗ

Група ІП-18-2

Чекін Олександр

2022

Міністерство освіти і науки України
Івано-Франківський національний технічний університет нафти і газу

Інститут інформаційних технологій

Кафедра інженерії програмного забезпечення

УДК 004.4

ДИПЛОМНА РОБОТА БАКАЛАВРА

Тема: «Інформаційні моделі та засоби роботи систем відео моніторингу»

(назва згідно з наказом ректора)

Спеціальність – 121 "Інженерія програмного забезпечення"

ПОЯСНЮВАЛЬНА ЗАПИСКА

ДРБ.ІП - 46.00.00.000 ПЗ

(позначення)

Студент

ІП-18-2 /Чекін О.І./
(шифр групи) (підпис) (дата) (розшифровка підпису)

Керівник роботи

доцент /Михайлюк І.Р./
(посада) (підпис) (дата) (розшифровка підпису)

Консультанти:

з питань нормоконтролю

доц., к.т.н /В.Я. Піх/
(посада) (підпис) (дата) (розшифровка підпису)

Допускається до захисту
Завідувач кафедри

проф., д.т.н. /В.І.Шекета/
(посада) (підпис) (дата) (розшифровка підпису)

2022
(рік)

Івано-Франківський національний технічний університет нафти і газу

Інститут, факультет інформаційних технологій

Кафедра інженерії програмного забезпечення

Ступінь вищої освіти бакалавр

Спеціальність 121 – Інженерія програмного забезпечення

ЗАТВЕРДЖУЮ:
Зав. кафедрою **ІПЗ**
проф. **В.І. Шекета**
“ ____ ” _____2022 р.

ЗАВДАННЯ

НА ДИПЛОМНУ РОБОТУ БАКАЛАВРА СТУДЕНТОВІ

Чекіну Олександру Івановичу

(прізвище, ім’я, по-батькові)

1. Тема проекту (роботи) «Інформаційні моделі та засоби роботи систем відео моніторингу» керівник проекту (роботи) Михайлюк Ірина Романівна, доцент

затверджені наказом закладу вищої освіти від **“** _ ” _____ 2022 р. № _____

2. Строк подання студентом проекту (роботи) **12 червня 2022 р.**

3. Вихідні дані до проекту (роботи) Результати і матеріали отримані під час проходження переддипломної практики

4. Зміст розрахунково - пояснювальної записки (перелік питань, які потрібно розробити)

1. Огляд існуючих концепцій, рішень та сервісів в даній області
2. Побудова моделі або алгоритму власного рішення
3. Документування реалізації власного оригінального рішення вибраними засобами

5. Перелік графічного матеріалу (з точним зазначенням обов’язкових креслень)

1. Загальна система відеоспостереження (рис. 1.1, ст. 10)
2 Загальна система відеоспостереження з кількома камерами (рис. 1.2, ст. 11)
3. Часова різниця (рис. 1.3, ст. 12)
4. Приклад віднімання фону для двох кадрів послідовності (рис. 1.4, ст. 13)
5. Загальна система віднімання фону (рис. 1.5, ст. 19)
6. Приклад віднімання фону (рис. 2.1, ст. 37)

6. Консультанти розділів проекту (роботи)

Розділ	Консультант	Підпис, дата	
		Завдання видав	Завдання прийняв

7. Дата видачі завдання _____

Керівник _____
(підпис)

Завдання прийняв до виконання_____
(підпис)

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів дипломного проекту (роботи)	Строк виконання етапів проекту	Примітка
1	Визначення та обгрунтування теми роботи	01.02.2022	виконано
2	Огляд існуючих концепцій, рішень та сервісів в даній області	01.03.2022	виконано
3	Побудова моделі або алгоритму власного рішення	01.04.2022	виконано
4	Документування реалізації власного оригінального рішення вибраними засобами	08.06.2022	виконано
6	Оформлення пояснювальної записки кваліфікаційної роботи	10.06.2022	виконано

Студент _____
(підпис)

Керівник роботи _____
(підпис)

АНОТАЦІЯ

В дипломній роботі виконано розробки програмного продукту виявлення в режимі реального часу рухомих та статичних середовищ спостереження, які контролюються статичними камерами. Це досягається на основі результатів, отриманих шляхом віднімання фону. Для цього завдання використовуються моделі гаусової суміші та оцінка щільності ядра. Ретельний огляд найсучасніших формулювань для використання у завданні фонового віднімання виявляє деякі можливості подальшого розвитку, які вирішуються в новому підході на основі, що включає схему контролю дисперсії. Запропонований метод підходу для параметричних і непараметричних даних дає нам кращий метод віднімання фону з більшою точністю та легшою параметризацією моделей для різних середовищ. Це також сходить до більш точних моделей сцен. Виявлення рухомих об'єктів досягається за допомогою результатів віднімання фону. Для виявлення нових статичних об'єктів використовуються дві фонові моделі, які навчаються з різною швидкістю.

КЛЮЧОВІ СЛОВА: РЕЖИМ РЕАЛЬНОГО ЧАСУ, МОНІТОРИНГ, ФОН, РУХОМИЙ ОБЄКТ, ПРОГРАМНИЙ ПРОДУКТ, ЯДРО.

ANNOTATION

In the thesis, the development of a software product for real-time detection of moving and static surveillance environments monitored by static cameras was carried out. This is achieved based on the results obtained by background subtraction. Gaussian mixture models and kernel density estimation are used for this task. A careful review of the state-of-the-art formulations for use in the background subtraction task reveals some opportunities for further development, which are addressed in a new approach based on incorporating a variance control scheme. The proposed approach method for parametric and non-parametric data gives us a better background subtraction method with higher accuracy and easier parameterization of models for different environments. This also comes down to more accurate scene models. Detection of moving objects is achieved using the results of background subtraction. To detect new static objects, two background models are used, which are trained at different speeds.

KEYWORDS: REAL-TIME MODE, MONITORING, BACKGROUND, MOVING OBJECT, SOFTWARE, KERNEL.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ 6
ВСТУП 7
1 ОГЛЯД ІСНУЮЧИХ ПРОГРАМНИХ РІШЕНЬ ТА СЕРВІСІВ В ПРЕДМЕТНІЙ ОБЛАСТІ 10
1.1 Системи відеоспостереження 10
1.2 Огляд структури роботи 15
1.3 Відповідні підходи 21
1.4 Сучасні тенденції та висновки 32
2 АЛГОРИТМІЧНІ ТА МАТЕМАТИЧНІ РІШЕННЯ ДЛЯ ЗАДАЧ ДАНОГО КЛАСУ 37
2.1 Віднімання фону 37
2.2 Алгоритм моделі гаусової суміші 40
2.3 Алгоритм оцінки щільності ядра 41
3 ВЛАСНЕ ОРИГІНАЛЬНЕ АЛГОРИТМІЧНЕ ТА ПРОГРАМНЕ РІШЕННЯ ЗАДАЧІ 45
3.1 Показники оцінки 45
3.2 Результати експерименту 46
3.3 Обговорення: 50
ВИСНОВКИ 52
СПИСОК ПОСИЛАНЬ НА ДЖЕРЕЛА 58
БІБЛІОГРАФІЧНА ДОВІДКА

ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ

РО – Рухомий об’єкт

БД – База даних

СУБД - Система управління базами даних

ПЗ - Програмне забезпечення

ВСТУП

Комп'ютерний зір, обробка зображень і розпізнавання образів – це величезні області досліджень, що стосуються автоматичного аналізу зображень і послідовностей, з широким спектром застосувань, таких як дистанційне зондування, медична діагностика, взаємодія людини з комп'ютером або стиснення відео, якщо згадати лише деякі. їх. Скориставшись досягненнями в цих галузях, роботизована відеорозвідка з'явилася як запит на те, що в останні роки привернуло багато уваги через збільшення небезпеки для відкритих місць безпеки, наприклад, залізничних станцій. або термінали літака. Суть полягає в тому, щоб допомогти людям-адміністраторам перевірити системи камер закритого телебачення, попередивши їх про відхилення від типової поведінки в регіоні, що спостерігається. Це дає фундаментальну перевагу в тому, що адміністратор може відображати більшу кількість камер, зосереджуючи увагу на основних фокусах у просторі та часі, в той час як структура очікує нудної перевірки доменів, де відбуваються події. Більше того, знання, отримані за допомогою методу для програмованих стратегій, можна використовувати як частину запиту, щоб допомогти адміністраторам відео та законним експертам відновити підтвердження із записаної відеоінформації, щоб регулювати розташування відео у великих зонах у завданнях, наприклад, панорамування та масштабування за допомогою камер, незважаючи на менш спеціалізовані питання, як-от забезпечення захисту людей.

Але використання систем відеоспостереження не обмежується додатками безпеки. Нині системи відеоспостереження також впроваджуються в універмагах, щоб забезпечити оцінку реклами та якості обслуговування, на автомагістралях з метою моніторингу дорожнього руху і навіть у будинках для людей похилого віку, щоб допомогти їм у неінвазивний спосіб. Цей успіх був підкріплений зниженням цін у галузі сенсорів, що здатне забезпечити більш якісні камери все менших розмірів за низькими цінами. Крім того, впровадження бездротових мереж означало різке зниження витрат на розгортання. З переходом

на мережі IP-камер, великі мережі камер можуть бути як локальними, так і дистанційно керованими.

Швидка розробка фреймворків відеорозвідки призводить до збільшення кількості відео, які слід перевіряти та зберігати в диспетчерській. Ці результати в постійно зростаючому робочому навантаженні для адміністраторів системи відеоспостереження, які перевантажені розташуваннями камер. Щоб полегшити цю проблему, запрограмовані процедури відеоекспертизи стосуються діяльності з розуміння та практик у відеопослідовності з конкретною кінцевою метою, щоб попередити адміністраторів відеоспостереження у разі певних обставин. Ця ситуація стосується проактивної сторони очікування протиправних дій. Крім того, системи відеоспостереження можуть також використовуватися для розслідування злочинів та притягнення до відповідальності правопорушників. Для ефективного виконання останнього завдання можна використовувати індексування та узагальнення відео. Крім того, автоматизовані системи відеоспостереження породили парадигму доведення інтелекту до мережі. Це дозволяє проектувати розподілені мережі спостереження, які потребують меншої пропускної здатності для передачі захопленої інформації.

Тим не менш, оскільки системи відеоспостереження стали повсюдно поширеними, деякі аспекти розгорнутих систем були піддані сумніву. Одним із аспектів є ефективність щодо запобігання злочинності. Іншим є потреба в захисті конфіденційності та безпеки особистої інформації, яка в останні роки привертає все більше уваги

Спеціалізоване створення та розсилання відеоспостережень включає різні ключові питання, які необхідно вирішити. Самий мінімальний рівень конфігурації фреймворка стосується проблем обладнання, включаючи отримання відео (камери), накопичення гаджетів і систем. На цьому рівні вибираються такі параметри, як топологія системи та відповідність. На цьому рівні дані, накопичені камерами, досліджуються методом для систем обробки зображень і відео, щоб виділити цінні дані з відеопослідовності. Це рівень, що

надає семантичні можливості фреймворку. Нарешті, на верхньому рівні витягнута інформація представляється користувачеві і в кінцевому підсумку зберігається в базі даних для подальшого використання. На цьому рівні слід враховувати міркування щодо ергономіки системи в цілому та взаємодії людини з комп’ютером. Очевидно, що рішення, прийняті на різних рівнях проектування, можуть вплинути на рішення, які будуть прийняті на інших рівнях; навіть більше у випадку внесення даних до мережі.

1 ОГЛЯД ІСНУЮЧИХ ПРОГРАМНИХ РІШЕНЬ ТА СЕРВІСІВ В ПРЕДМЕТНІЙ ОБЛАСТІ

1.1 Системи відеоспостереження

Автоматизовані системи відеоспостереження, які в цій роботі для стислості називаються системами спостереження (інакше чітко зазначено), покладаються на автоматичне виявлення подій, що представляють інтерес, за допомогою кількох методів аналізу, які в основному випливають із комп'ютерного зору, обробки зображень та розпізнавання образів. Виявлення подій, що представляють інтерес, є завданням, що залежить від програми, і до нього можна підходити дуже різними способами. Тим не менш, існує загальна кількість кроків, через які зазвичай проходить загальна система спостереження, а саме виявлення об'єктів, об'єднання об'єктів, які зазвичай називають відстеженням, і розуміння сцени, що часто виконується менш амбітним завданням виявлення подій. Для успішного виконання цих завдань камери мають бути відкалібровані щодо зовнішнього декартового опорного простору, що дозволяє вимірювати розмір і положення виявлених об'єктів. Ці основні будівельні блоки автоматизованої системи відеоспостереження зображені на рисунку 1.1 і коротко представлені в наступних підрозділах.

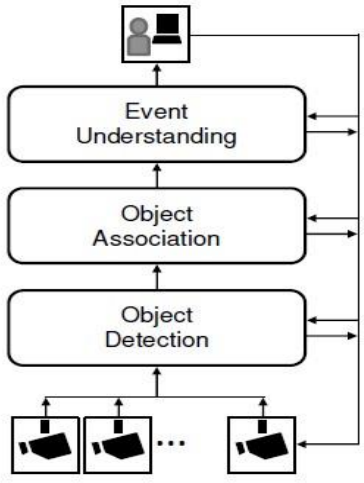


Рисунок 1.1 - Загальна система відеоспостереження

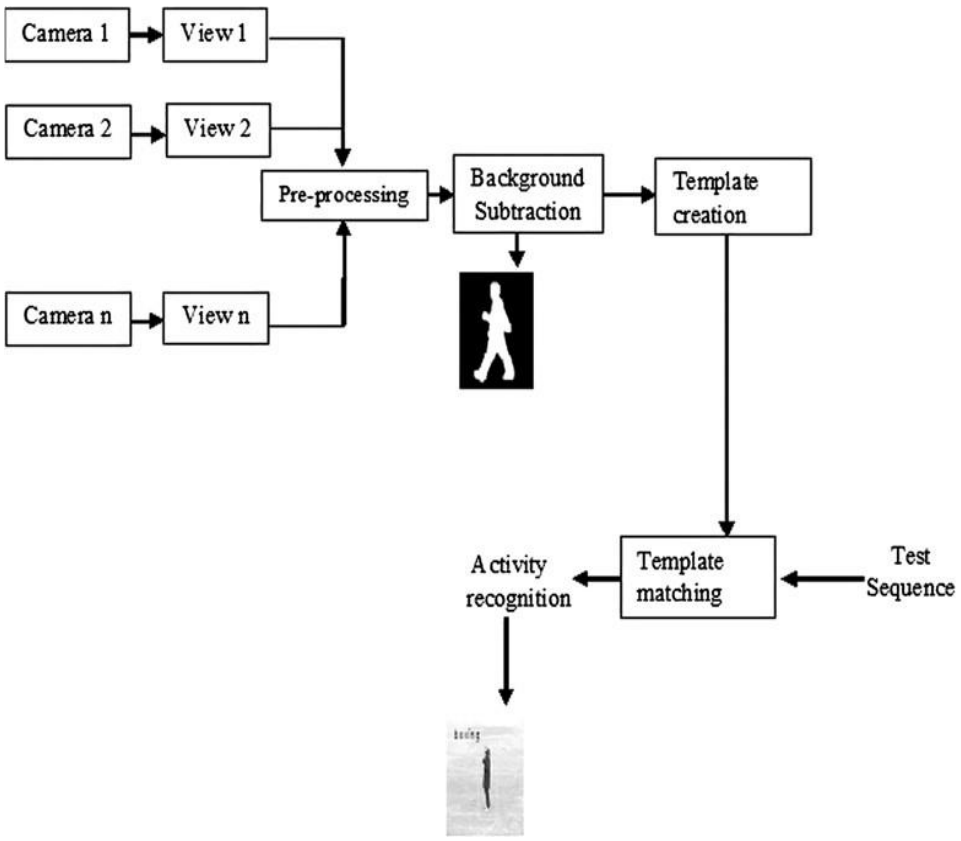


Рисунок 1.2 - Загальна система відеоспостереження з кількома камерами

Виявлення та класифікація об'єктів

Загальне розпізнавання об’єктів, також відоме як розпізнавання об’єктів на рівні категорії, вважається однією з найскладніших візуальних завдань комп’ютерного зору [9]. З огляду на будь-який екземпляр певного загального класу, наприклад, «людина», «автомобіль» або «велосипед», завдання полягає в тому, щоб правильно локалізувати та класифікувати його за допомогою візуальних ознак. Ретельна переслідування всіх моделей протесту та областей зображення може бути надто виснажливим для деяких програм комп’ютерного зору. Щоб зменшити багатогранність питання, розвідувальні рамки здебільшого розбивають питання на два етапи: для початку розпізнаються об’єкти інтриги, а по-друге, упорядковуються визначні статті. Об’єкти інтриги зазвичай характеризуються як статті, що представляють певний прогрес у спостережуваній сцені та загалом пов’язані з рухомими предметами.

До виявлення об’єктів можна підійти за допомогою трьох різних методів: тимчасове диференціювання, віднімання фону та оптичний потік. Ці три методи дають низькорівневий порядок пікселів. Щоб зібрати об’єкти, пікселі потім групуються, піклуючись про це розташування та їх просторове налаштування. Світові відмінності залежать від визначення відмінностей між обрисами відео в кожній позиції пікселя та визначення як змінених пікселів тих, контрастність яких перевищує задану межу. Коротка диференціація надзвичайно універсальна для ситуацій елементів і низька вимогливість у обчислювальних термінах, однак вона нехтує розділенням усього розташування пікселів у порівнянні з артикулами, що рухаються. Ранні роботи, засновані на часовій диференціації, можна знайти в [9] та посилання в ній. Віднімання фону є найбільш поширеним підходом у налаштуваннях зі статичними камерами. Він полягає у використанні моделі фону сцени для виявлення об’єктів переднього плану, розрізняючи вхідні кадри з моделлю. Фонове віднімання в основному швидке та має низькі обчислювальні вимоги. Однак він може бути чутливим до різких змін освітлення та невеликих рухів камери, наприклад, до вібрацій. Хороший вступ до фонового віднімання, включаючи основні проблеми, з якими має працювати підхід до фонового віднімання, можна знайти в [8]. Оптичний потік — це оцінка, яка використовується для визначення відповідних точок між двома зображеннями. Методи на основі оптичного потоку можна використовувати для виявлення незалежно рухомих об’єктів навіть за наявності руху камери. Тим не менш, навіть у своїх найефективніших реалізаціях вони дуже вимогливі з точки зору обчислень. Крім того, залежно від обмеження плавності, відповідні точки в розглянутих кадрах не можуть бути віддалені більше ніж на кілька пікселів, отже, швидкість руху об’єктів і камери обмежена. Хороший вступ до теми обчислення оптичного потоку можна знайти в [1]. З оглядом найсучасніших підходів та їхньої ефективності можна ознайомитися в режимі онлайн на веб-сайті набору даних Middlebury2 .



Рисунок 1.3 - Часова різниця

Зліва направо: перше зображення пари, що містить одну рухому особу, друге зображення тієї ж пари зображень і маску відмінності.



Рисунок 1.4 - Приклад віднімання фону для двох кадрів послідовності

Відстеження об'єктів/відстеження руху

Відстеження об’єкта – це завдання встановлення відповідності між виділеними протестами над корпусами відеоряду. З конкретною кінцевою метою для досягнення цієї мети використовується модель статей і рух, який вони демонструють. Звичайними моделями запитань є фокуси, примітивні геометричні фігури, як, наприклад, овали та прямокутники, контури, моделі пояснюваних фігур і скелети. Залежно від використаного вибраного питання, модель руху може бути розмежована. Наприклад, якщо об’єкт представлений

точкою, то можна використовувати лише трансляційну модель, тоді як у випадку більш складних моделей об’єктів, наприклад, силуетів, можна використовувати параметричні та непараметричні моделі руху. Залежно від домену програми робляться припущення, щоб обмежити проблему відстеження. У сфері спостереження точкові моделі відстеження є популярним вибором для вирішення проблеми відстеження. Таким чином, зазвичай є методами оцінки стану, що використовуються для обчислення вартості даної асоціації об’єктів. Чудовий вступ до теми відстеження та важливих пов’язаних питань, включаючи використання відповідних функцій зображення, вибір моделей руху та виявлення об’єктів, можна знайти.

Виявлення об'єктів можна здійснити багатьма способами. Як це зробити, залежить від наявних даних і від того, рухається об’єкт чи ні. Для об’єктів у стані спокою необхідно знати певні попередні знання про тип об’єктів. Це може бути один зразок зображення об’єкта для відстеження. Виявлення рухомих об’єктів у послідовності зображень не потребує попередніх знань, але потребує кількох послідовних зображень. Два поширені методи виявлення рухомих об'єктів [8]:

Віднімання фону

Віднімання фону – це широко використовуваний підхід для виявлення рухомих об’єктів із статичних камер. Фундаментальна логіка полягає в тому, щоб виявити об'єкти з різниці між поточним кадром і опорним кадром, який називається фоновим зображенням. Принцип полягає в тому, що якщо відоме еталонне фонове зображення, це зображення можна порівняти з кадром, в якому об’єкти мають бути виявлені.

Оптичний потік

Обчислюючи поле потоку пікселів у послідовних кадрах, можна виявити об’єкти.

Кластери пікселів, що рухаються разом, імовірно, є частиною одного об’єкта.

Коли місце розташування об’єкта, який слід відстежувати, відомо, деякі ознаки повинні бути виділені та записані, щоб зробити можливим знайти той самий об’єкт у нових кадрах. Хороша сегментація на фоні гарантує, що будуть записані лише об’єкти, які насправді належать об’єкту, що цікавить. Таким чином, проблема полягає в тому, щоб визначити, які пікселі належать до об’єкта, а які — до фону, якщо дана область, що містить об’єкт. У деяких випадках піксельна сегментація не потрібна, але якщо в об’єктну модель буде включено занадто багато фону, шум буде дуже важко відстежувати ціль.

Об’єкти можуть бути представлені кількома способами, як центроїд, кілька точок, примітивні геометричні фігури або контури та силуети об’єктів. Їх можна об’єднати, щоб отримати гарне уявлення об’єкта, який слід відстежувати. Добрими функціями для відстеження є речі, які продовжують виглядати так само, навіть якщо масштаб змінюється або об’єкт повертається за межі площини. Прикладами такого роду ознак є кути та краї. Іншим можливим уявленням об'єкта є кольорова гістограма області об'єкта.

Віднімання фону

У найфундаментальнішому сенсі віднімання фону — це саме те, що завершує назва, загальний контраст між еталонною картинкою (фоном) і картинкою, що цікавить. У позиціях зображення, де відмінність є більш помітною, ніж деякий край, позиція класифікується, не маючи місця з основою, тобто називається переднім пікселем. Більшість сучасних обчислень для виконання віднімання фундаменту є більш приголомшливими, ніж це, і їх можна розділити на кілька класифікацій. Принциповий контраст між більшістю стратегій – це засоби, за допомогою яких розмовляють з базовою моделлю. Від легких до більш складних:

Для кожного пікселя фон демонструється незалежно як функція щільності ймовірності Гаусса. Гауссове апропріація підбирається до n останніх значень пікселів, і піксель упорядковується шляхом встановлення ймовірності того, що останній піксель відображає питання, яке не можна відрізнити від попередніх значень пікселів.

Іноді частина зображення, яку слід класифікувати як фон, не є повністю статичною, деякі частини можуть трохи рухатися (через вітер, вібрації камери тощо) і все одно їх слід класифікувати як фон. Для адаптації до такого типу фону єдиної цінної моделі фону недостатньо. Ідея полягає в тому, щоб мати відмінні моделі Гаусса для різних можливих фонових об’єктів, якщо оцінка пікселя, ймовірно, не походить від будь-якого з різноманітних засобів передачі, тоді вона називається переднім планом.

Оцінка щільності ядра (KDE)

У цьому методі будується функція, яка дає ймовірність того, що даний піксель належить до розподілу фонових пікселів. Для гауссового бігучого середнього попередні відомі значення пікселя були підігнані до гауссового значення для моделювання розподілу, в оцінювачі щільності ядра розподіл замість цього будується із суми ядер.

1.2 Огляд структури роботи

У центрі уваги даної дипломної роботи є виявлення об'єктів у необмежених середовищах, які контролюються відеокамерами. Об'єктами інтриги є рухомі й на додаток нові статичні статті. Структура відеорозслідування не містить жодної попередньої інформації ні про спостережувану сцену, ні про зовнішній вигляд статей, які потрібно розпізнати. Основним підходом у верхній частині списку пріоритетів створених алгоритмів є розміщення покинутих об’єктів на відкритих просторах, що набуло важливої уваги в зоні безпеки, оскільки здані предмети можуть регулярно розглядатися як небезпека для людей. загальна безпека

населення. Остаточна система має надавати онлайн-попередження операторам. Крім того, виявлені рухомі об’єкти повинні бути надані інструментам аналізу вищого рівня, щоб розпізнати подальші дії та поведінку, що представляють інтерес, типові для систем спостереження для громадських приміщень.

Тоді віднімання фону буде зроблено алгоритмами для відстеження об’єкта. Для фонового віднімання використовуються такі методи:

Розширена модель суміші Гаусса (GMM) для додатків відеоспостереження, яка включає останні пропозиції щодо покращення продуктивності системи та конвергенції системи, а також нову евристику для:

– Краща ініціалізація параметрів нових створених режимів і – Уникнення появи наддомінуючих режимів.

• Метод оцінки щільності ядра (KDE) будується функція, яка дає ймовірність того, що даний піксель належить до розподілу фонових пікселів. Для гауссового бігучого середнього попередні відомі значення пікселя були підігнані до гауссового значення для моделювання розподілу, в оцінювачі щільності ядра розподіл замість цього будується із суми ядер.

• У подальшому відеоінженерія виконана, за допомогою алгоритму віднімання фону, фон з одного відео та передній план з іншого відео буде відніматися, і ми об’єднаємо їх в одне відео. Таким чином, ми також можемо зробити більше відеоінженерії з різними відео.

Виявлення змін – це завдання низького рівня зору, яке використовується як початкова фаза в багатьох програмах комп’ютерного зору, наприклад, відеоспостереження, низькошвидкісне кодування відео, підключення людини до комп’ютера, доповнена реальність або медичні знахідки, і це лише кілька з них. З огляду на групування зображень, мета полягає в тому, щоб розрізнити для кожного кадру розташування пікселів, які принципово не зовсім такі самі, як попередні краї. Залежно від застосування, потреби та імперативи розрахунку відкриття є відмінними. Аналогічно, значення того, що по суті є унікальним, може залежати від домену програми.

У сфері відеоспостереження виявлення змін регулярно використовується як частина запиту на об’єкти переднього плану з фону. Об'єкти переднього плану - це об'єкти автоматизованої системи спостереження. Розділені об’єкти переднього плану потім пов’язуються між кадрами з конкретною кінцевою метою, щоб розіграти розслідування сцени та визначити випадки премії. Таким чином, прийнято, що фон можна відобразити за допомогою методу для статистичної моделі, моделі фону. У будь-якому випадку, є деякі фонові характеристики, такі як рухоме листя або раптові зміни освітлення, які можуть ускладнити завдання демонстрації та догляду за фундаментом. Всебічне дослідження основних проблем і деяких принципів, які можуть бути використані для їх вирішення, можна знайти в [8]. Сегментація об’єктів переднього плану за допомогою виявлення змін щодо моделі фону зазвичай відома як віднімання фону. На рисунку 1.5 зображена базова схема загальної системи віднімання фону. Основними проблемами, з якими має впоратися алгоритм фонового віднімання, є:

• Поступові зміни освітлення, які в основному спостерігаються на відкритому повітрі в різний час доби і впливають на зовнішній вигляд об'єктів на спостережуваній сцені.

• Раптові зміни освітленості, які в основному виникають у приміщеннях через увімкнення та вимкнення штучних джерел світла, а на вулиці – через нестабільні погодні умови, коли хмари раптово приховують сонце.

• Тіні, які в основному відкидаються рухомими об'єктами і ускладнюють точну сегментацію об'єктів (статичні об'єкти, що належать до фону, також відкидають тіні; проте це не так вже й проблематично для процесу віднімання фону, оскільки вони завжди відкидаються в одній позиції - або в повільних позиціях у сценарії на відкритому повітрі, залежно від положення сонця - і може бути легше розміщений у фоновій моделі).

• Динамічний фон, тобто ті частини фону, які мають різний вигляд через те, що містять якісь рухомі об'єкти, як-от дерева, що хвилюються, брижі води, ескалатори тощо, які більше не представляють інтересу для інтерпретації сцени.

• Камуфляж, створений об'єктами, зовнішній вигляд яких важко відрізнити від зовнішнього вигляду фону.

• Bootstrapping, який необхідний через загальну недоцільність навчання фонової моделі з абсолютно порожньою сценою.

Насправді, автори також вказали на деякі проблеми, які, на їхню думку, повинна впоратися система фонового обслуговування:

• Переміщені об'єкти, що відноситься до виявлення, що відповідають фоновим об'єктам, які були переміщені.

• Спляча людина, що стосується об'єктів переднього плану, які з'являються на сцені та залишаються нерухомими через деякий час.

• Людина, що ходить, що стосується об’єктів, які були вивчені як частина фону і в певний момент часу починають рухатися та залишають місце події.

Тим не менш, ці три труднощі не розглядалися в цій роботі як природні для проблеми віднімання фону, оскільки ці питання слід розглядати відповідно до застосування, яке знаходиться у верхній частині списку пріоритетів. Фактично, момент часу, з якого, наприклад, чоловік, який киває головою, більше не є захоплюючим, слід характеризувати даною заявкою і, таким чином, не слід розглядати як загальну фонову проблему утримання. Крім того, дивно, що ці три проблеми також можна розглядати як три особливості, з якими повинна впоратися гідна техніка завантаження. Нарешті, проблема розриву області переднього плану, також зазначена, яка включає в себе неможливість розпізнавання внутрішніх пікселів питання через однорідність затінення, не розглядалася в цій роботі як загальна проблема виявлення змін, оскільки здебільшого стосується методологій, заснованих на накресленні відмінностей.



Рисунок 1.5 - Загальна система віднімання фону.

Таксономія

Підходи до фонового віднімання можна розділити на рекурсивні та нерекурсивні. Рекурсивні підходи оновлюють фонову модель у міру надходження нових спостережень, тому споживають низькі ресурси з точки зору вимог до обчислень і пам'яті. З іншого боку, нерекурсивні підходи зберігають буфер останніх вхідних відеокадрів для оцінки фону. Тому нерекурсивні підходи мають більш високі вимоги до пам'яті. Тим не менш, оскільки вони мають копію останніх відеокадрів, вони можуть впоратися з деякими проблемами, такими як відхилення викидів і швидка конвергенція, з якими не можна легко впоратися за допомогою рекурсивних методів.

Нерекурсівні Техніки

Нерекурсивний підхід використовує ідею розсувного вікна для віднімання основи. Він закріплює/буферує прохідний відеокорпус і вимірює основу з огляду на світову різноманітність кожного пікселя з підтримкою. Для цієї ситуації передумова ємності висока, ці процедури дуже універсальні. Щоб зрозуміти

проблему накопичення, ми можемо зберігати контури з помірною швидкістю обсадки. Нижче наведено кілька стратегій для систем.

Рамка диференція

Диференціація кадрів використовує відеокадр у момент часу, t-1, як фонову модель для кадру в момент часу t. Оскільки він використовує лише один попередній кадр, розрізнення кадрів може не виявити внутрішній піксель великого, рівномірно забарвленого рухомого об’єкта.

Медіанна фільтрація

Це найбільш поширена техніка для формування фону. Оцінка фону визначається як медіана для кожного пікселя всіх кадрів у буфері, припущення полягає в тому, що піксель залишається у фоновому режимі більше половини кадрів у буфері . Медіанну фільтрацію було розширено на колір, замінивши медіану на Середній.

Фільтр лінійного прогнозу

Він обчислює поточну оцінку фону, застосовуючи фільтр лінійного прогнозування до пікселів у буфері. Коефіцієнти фільтра оцінюються під час кожного кадру на основі коваріацій вибірки, що ускладнює застосування цього методу в реальному часі.

Рекурсийна техніка

Рекурсивні методи не підтримують віднімання фону. Швидше, вони рекурсивно переглядають одиночний фон, який демонструють з огляду на кожен вхідний кадр. Тому контури введення з видаленого минулого можуть вплинути на поточне фонове відображення. За умови, що ми порівняли нерекурсивний метод, ця система вимагає меншої потужності, однак будь-яка помилка моделі,

яка не помічається, може тривати тривалий проміжок часу. Нижче описано кілька стратегій використання методів.

Наключений метод фільтра

Цей метод був використаний у фоновому моделюванні для моніторингу міського руху . У цій схемі поточна оцінка медіани збільшується на одиницю, якщо вхідний піксель більший за оцінку, і зменшується на один, якщо менше. Ця оцінка в кінцевому підсумку сходить до значення, для якого половина вхідних пікселів більша, а половина менша за це значення, тобто медіана.

Суміш гауса (MoG)

Цей метод відстежує декілька розподілів Гаусса, MoG користується величезною популярністю з моменту його вперше запропонованого. Цей метод підтримує функцію щільності для кожного пікселя. Таким чином, він здатний обробляти віднімання фону кількох моделей.

1.3 Відповідні підходи

У цьому методі різниця обчислюється між двома кадрами, з яких один є поточним кадром, а інший — фоновим кадром, щоб виявити наявність будь-якого рухомого об’єкта у відео. Рівняння для цього

$$|\text{рамка } Ic - \text{рамка } Ib| > T$$

У цьому кадрі Ic — поточний кадр, кадр Ib — фоновий кадр, а T — порогове значення.

Для цього кроки Алгоритму такі:

• Визначте фоновий кадр і поточний кадр із відеопотоку.

• Обчисліть перетворене зображення цих кадрів у шкалі сірого.

• Зафіксуйте розмір кадру для подальшого обчислення пікселів.

• Обчисліть різницю між пікселями двох кадрів і звірте з визначеним пороговим значенням.

• Якщо різниця перевищує порогове значення, візьміть її як об'єкт переднього плану, інакше як фон.

• Оновити порогове значення відповідно до змін у послідовних кадрах.

Переваги використання цього методу полягають у тому, що його швидко простіше застосовувати і добре працює для статичного фону, але йому потрібен фон без об’єктів, інакше цим методом їх можна вважати рухомим об’єктом.

Нижче наведено набір наукових робіт, пов'язаних з цією технікою:

Покращений алгоритм виявлення рухомого об'єкта, заснований на розрізненні кадрів і виявленні країв: виявлення рухомого об'єкта та віднімання є важкою роботою. Він представив підхід до виявлення рухомого об'єкта, а потім відняв його з кадру за допомогою методу диференціації кадрів. Перш за все, він виявляє краї кожного двох безперервних кадрів, а потім отримує різницю між двома краями зображення. А потім він ділить зображення різниці країв на кілька маленьких блоків і вирішує, рухаються вони чи стають, порівнюючи кількість ненульових пікселів.

Було помічено, що покращений алгоритм визначення рухомого об’єкта та алгоритм віднімання, заснований на диференціації кадрів, має набагато більшу швидкість розпізнавання та більшу швидкість виявлення, ніж кілька класичні алгоритми. Цей алгоритм буде здаватися окремим хибним у складнішому фоні, і ще є куди вдосконалюватися.

Вилучення відеооб’єктів на основі DFD між кадровою та пороговою сегментацією: розглядає проблему складного руху та неохопленого фону в сегментації фону, був запропонований новий метод на основі DFD між кадрами та пороговою сегментацією. У цьому методі відфільтровано і отримано дві послідовні різниці між кадрами, а потім змінено різні зображення за допомогою «заповненої асиміляції», щоб отримати шаблон відмінності та використовувати буфер шаблону для підтримки цілісності шаблону ітерації. Цей алгоритм не

залежить від фіксованого фону і може усунути непокритий фон у відмінних зображеннях.

Було видно, що результат вилучення відеооб'єкта для однієї рухомої цільової відеопослідовності є задовільним. І він може ефективно подолати шум, єдину мету руху складності та вплив фонової експозиції при розділенні відеооб’єкта зі зміною виявлення.

Відстеження об’єктів за допомогою розрізнення кадрів і відповідності шаблону: систему відстеження та вилучення об’єктів з використанням розрізнення кадрів і відповідності шаблону. Кадрова диференціація використовується покадрово для ефективного виявлення рухомого об’єкта. Зображення шаблону використовується для відповідності та генерується динамічно, що гарантує, що зміна орієнтації та положення об’єкта не впливає на систему. Автор посилається на споріднену роботу де було помічено, що цей метод був високоефективним і може використовуватися як інструмент спостереження в різних програмах. Цей метод також забезпечує кращі результати для вилучення об’єктів, які можна легко застосувати до низки полів. Цей метод також можна використовувати для вилучення об’єкта, який знаходиться у віддаленій точці. У майбутньому, щоб підвищити ефективність, можна буде зробити більше над цим.

BSFD: Алгоритм розрізнення кадрів фонового віднімання для виявлення та вилучення рухомих об’єктів: представляє два поширені алгоритми визначення рухомого об’єкта, віднімання фону та розрізнення кадрів, а також їх порівняння. Фонове зображення, яке використовується для обробки наступного кадру, створюється через суперпозицію поточного кадру. Цей алгоритм робить об’єкт, який довго зберігається, але не виявляється як частина фону.

Було помічено, що алгоритм може виявляти рухомий об'єкт більш ефективно і точно. Він усунув недоліки методу віднімання фону та методу різниці кадрів, запропонував динамічне оновлення фонового зображення методом розрізнення кадрів і використовує потужність методу віднімання фону.

Вилучення рухомих об’єктів за допомогою розрізнення кадрів, видалення привидів і тіней. Тут представлена техніка для вилучення рухомих об’єктів на основі тимчасової диференціації, видалення привидів за допомогою NCC, під час використання камери з нестатичним панорамним нахилом. Щоб виявити рухомий об’єкт на поточному зображенні, попередній кадр зображення, ft-1, компенсується відносно поточного зображення. Тут пропонується метод видалення за допомогою попереднього кадру зображення, фут-2. Потім вихід очищається за допомогою оператора морфологічного відкриття перед видаленням тіні. Кожні попередньо визначені пікселі переднього плану класифікуються на пікселі тіні або пікселі фону.

Було помічено, що цей метод показує, що рухомі об'єкти вилучаються без тіней. Цей метод можна використовувати в режимі реального часу з високою швидкістю обчислень і відмінною продуктивністю при виявленні рухомого об’єкта в кожному кадрі. Існує багато додатків, які використовують цю систему, наприклад, система спостереження в житловій зоні, відстеження людей і дорожній рух. Майбутня робота для цього проекту полягає в подальшому покращенні виявлення тіні з тонкою формою від рухомих об’єктів.

Таблиця 1.1 - Підсумок різниці кадрів

Рік	Автор	Назва	Опис
2007	Zhan Chaohui	Покращене виявлення рухомих об'єктів алгоритм, заснований на розрізненні кадрів і виявленні країв.	Виявити проблему віднімання фону в диференціації кадрів і надати вдосконалений метод вирішення проблема з високою швидкістю виявлення та вирішення складної фонової проблеми.
2008	Jinwei Cui	Вилучення відеооб'єктів на основі DFD між кадровою та пороговою сегментацією.	У цьому методі автор усуває складний рух і незакритий фон і пропонує новий метод DFD.

2012	N. Prabhakar	Відстеження об'єкта за допомогою рамки розрізнення та відповідність шаблону.	Розрізнення кадрів і відповідність шаблону використовуються для виявлення об’єкта та ефективного його вилучення. Це Метод є високоефективним і може використовуватися як інструмент спостереження в різних програмах.
2014	D. Stalin Alex	BSFD: Фонова диференціація кадрів віднімання алгоритм для виявлення та вилучення рухомих об'єктів.	У цьому автор порівнює два алгоритми віднімання об'єктів. Виправлено їх недоліки та запропоновано динамічно оновлюваний метод.

Модель гаусової суміші (GMM)

Алгоритм моделі гаусової суміші (GMM) заснований на припущенні, що фон видимий частіше, ніж передній план, а дисперсія фону невелика. Оскільки один гаусів не є пристойною моделлю для сцен на вулиці, цей метод віднімання фону був запропонований Штауффером і Грімсоном, в якому кожен піксель у фоновому режимі моделюється як суміш Гаусса. Кожне значення пікселя узгоджується з поточним набором моделей, щоб знайти відповідність. Якщо збігу не знайдено, найменша отримана модель відхиляється, і вона замінюється на новий гауссовий з ініціалізацією наявним значенням пікселя, що означає, що значення пікселя, яке не відповідає фону, вважається фоном. Цей метод вимагає менше пам’яті для роботи і дає дуже точні результати, а також може мати справу з повільними змінами освітлення, хоча він не може впоратися з мультимодальним фоном і вимагає суворих обчислень.

Розуміння моделі фонової суміші для віднімання фону: представив основну теорію для розуміння базової моделі та навчання шляхом реалізації алгоритму Штауффера-Грімсона з різними параметрами. Це в основному показує, які наближення до теорії були зроблені та як покращити стандартний алгоритм шляхом перевизначення цих наближень.

У ньому перераховані всі основні параметри моделі та типові значення, а також розширення, які необхідні для практичного використання алгоритму. Ця

робота надала теоретичний інструмент, за допомогою якого можна було змінити або адаптувати оригінальний алгоритм для кращої продуктивності, більшої швидкості та надання інформації, необхідної для швидкої реалізації.

У ньому було сказано, що віднімання фону є важливою обробкою для багатьох відеододатків. Байєсівське формулювання фонової сегментації на основі моделі Гаусової суміші. Вони показують, що задача складається з двох задач оцінки щільності, одна не залежить від застосування, а інша залежить від застосування, і для обох можна отримати набір теоретично оптимального рішення. Ця робота була перевірена на відео зустрічей та відео про дорожній рух.

Показано, що набір інтуїтивно зрозумілих і теоретично обґрунтованих рішень можна сформулювати в термінах задачі оцінки щільності. За допомогою цього запропонованого алгоритму рішення цих проблем, фреймворк було застосовано для сегментації відеозаписів зустрічей і трафіку. Продуктивність над існуючим методом підтверджує цю теорію.

Покращена модель адаптивної гауссової суміші для віднімання фону: віднімання фону є комп'ютерним завданням комп'ютерного зору. Це звичайний підхід на рівні пікселів. У цьому ефективному адаптивному алгоритмі з використанням щільності ймовірності суміші Гаусса було розроблено рекурсивне рівняння, яке постійно використовувалося для оновлення параметрів, а також одночасного вибору відповідної кількості компонентів для кожного пікселя.

Тут була представлена вдосконалена схема віднімання фону GMM. Цей новий алгоритм може автоматично вибирати необхідну кількість компонентів на піксель і таким чином повністю адаптуватися до спостережуваної сцени. При цьому час обробки зменшується, а сегментація також покращується.

Покращений алгоритм адаптивного моделювання фону, заснований на моделі суміші Гаусса: одну з найкращих моделей GMM для віднімання фонової сцени з повторюваним рухом. До цієї проблеми було запропоновано безліч підходів, які відрізняються типом фонової моделі, але він був одним із

найкращих. Однак велика кількість обчислень обмежила його застосування. Більше того, він мав труднощі з сегментацією об’єктів, що повільно рухаються, та об’єктів, які на деякий час зупиняються під час руху. На основі GMM (Gaussian Mixture Model) в алгоритмі був використаний адаптивний метод для зменшення обсягу обчислень, а також запропоновано адаптаційний метод з адаптацією швидкості навчання, щоб точно сегментувати об’єкти, які рухаються повільно або на деякий час зупиняються.

Було помічено, що порівняння запропонованого алгоритму з методом GMM мало чимало відмінностей. Результати сегментації показують, що запропонований метод мав кращу продуктивність, ніж метод GMM.

Адаптивний підхід GMM до віднімання фону для застосування в режимі спостереження в реальному часі: У цій новій моделі для віднімання фону в реальному часі за допомогою GMM (Adaptive Gaussian Mixture Model) було запропоновано. Цей новий метод був універсальним і адаптованим до динамічного фону, швидке освітлення змінює повторювані рухи. Тут також був вбудований метод виявлення тіні за допомогою колірної моделі horpresert. Цей метод можна використовувати для моніторингу районів, де доступ до руху дуже обмежений. Таким чином, при виявленні будь-яких несподіваних подій на сцені може бути спрацьована тривога, і ми можемо досягти спостереження в реальному часі навіть за відсутності постійного контролю людини.

Результати цього фонового віднімання (AGMM) є високоефективними. Цей метод можна використовувати для виявлення залишеного багажу в аеропорту та на вокзалах у будь-якому місці, де безпека є першочерговою проблемою. Цей метод може бути реалізований так, що будь-який рух у зоні може бути негайно виявлено і може бути викликано тривогу.

Новий метод виявлення рухомого об’єкта, заснований на покращеній різниці кадрів та покращеній моделі гаусової суміші: в існуючому методі виявлення руху, який включає віднімання фону та різницю кадрів. Але за допомогою методу різниці кадрів часто виникають деякі діри, і важко

побудувати фонову модель за допомогою методу віднімання фону. Отже, попередній алгоритм не досяг ідеальних результатів. Основною метою автора є поєднання методу різниці кадрів, покращеного за допомогою зображення історії руху, з методом віднімання фону на основі вдосконаленої моделі гаусової суміші для виявлення рухомого об’єкта. Було помічено, що покращена різниця кадрів використовувалася для виявлення об'єкта руху в часовій області, а покращене віднімання фону використовувалося для виявлення об'єкта руху в просторовій області. Нарешті, щоб частина були об'єднані, щоб отримати повний рух об'єкта. Цей алгоритм обробив багато відео та отримав задовільні результати.

Таблиця 1.2 - Резюме моделі гаусової суміші

Рік	Автор	Назва	Опис
2002	P. Wayne Power	Розуміння моделі змішування фону для віднімання фону.	Цей метод в основному показує, які наближення до теорії були зроблені і як покращити стандартний алгоритм шляхом перевизначення цих наближень.
2003	Dar-Shyang Lee	Байєсівська структура для моделювання фону суміші Гаусса.	Байєсівська формулювання фонової сегментації на основі моделі гаусової суміші. Це також свідчить про те, що проблема складається з двох задач оцінки щільності, одна не залежить від застосування, а інша залежить, і рішення також можна отримати для обох.
2004	Zoran Zivkovic	Покращена адаптивність Гаусова суміш Модель фонового віднімання.	Розроблено ефективний адаптивний алгоритм із використанням щільності ймовірності суміші Гаусса. Час обробки зменшується, а сегментація також покращується.
2008	Peng Suo	Покращений адаптивний фон Алгоритм моделювання на основі моделі гаусової суміші.	У цій моделі гаусової суміші можна відняти фонову сцену з повторюваним рухом. Адаптивний метод був використаний, щоб зменшити обсяг обчислень і точно сегментувати об’єкт, який рухається повільно або зупиняється.

2013	Subra Mukherjee*et al	Адаптивний підхід GMM до віднімання фону для застосування в режимі реального часу спостереження.	Було запропоновано новий підхід, який є надійним і швидко адаптованим до динамічного фону зміна освітлення, повторювані рухи. Цей метод може бути реалізований таким чином, що будь-який рух у зоні може бути негайно виявлено та може бути викликано тривогу.
2013	Ю Сяоян	Нове виявлення рухомого об'єкта	Основною метою цього є поєднання методу різниці кадрів, покращеного за допомогою рухомого зображення, з методом віднімання фону на основі покращеної моделі гаусової суміші для виявлення рухомого об’єкта. Цей алгоритм обробив багато відео та отримав задовільні результати.

Метод апроксимованого медіанного фільтра

Запропонували простий рекурсивний фільтр для оцінки медіани пікселя зображення, в якому поточна оцінка медіани збільшується на одиницю, якщо вхідний піксель більше оцінки, і так далі зменшується на одиницю, якщо вхідний піксель менший за оцінку. Оцінка в кінцевому підсумку сходить до значення, для якого половина вхідного пікселя більша, а половина пікселів менша за це значення, тобто це значення є медіаною.

У цьому процесі середня фільтрація буферує попередні N кадрів відеопотоку. Після цього фоновий кадр обчислюється з медіани буферизованого кадру, а фон віднімається від поточного кадру, щоб отримати піксель переднього плану.

Недоліками цієї техніки є те, що вона не пропонує більш гладкі результати за будь-яких обставин, оскільки це рекурсивна техніка, вона не зберігає буфер для оцінки фону на своєму місці, вона регулярно оновлює один фоновий кадр, таким чином, будь-який вхідний кадр із дуже далекого минулого може вплинути на поточну фонову модель. Хоча це означає, що він вимагає менше пам’яті, оскільки не потребує підтримки буфера.

Сегментація рухомого транспортного засобу в динамічному фоні за допомогою самоадаптивного методу це вводить адаптивний фільтр Калмана для моделювання динамічного фону для віднімання фону. Метод віднімання фону використовується для ідентифікації об'єкта та відомого, що використовується при сегментації рухомих об'єктів. Це також досліджує порівняльне дослідження методу віднімання Гаусса, диференціації кадрів і методу наближеного медіан.

Було помічено, що з рівняння фільтра Калмана ми можемо досягти точного виявлення об'єкта. Крім того, сегмент покращився, а виявлення об’єктів стало більш гладким.

Таблиця 1.3 - Підсумок методу апроксимованого медіанного фільтра

Рік	Автор	Назва	Опис
2011	Автор К.А. Ахмад	Сегментація рухомого транспортного засобу в динаміці	Цей метод є новим для фонового віднімання, а також з
2014	Алок Кумар Сінгх Кушвахе	Комплексний вейвлет Переміщення на основі об'єкт сегментація за допомогою на основі приблизного медіанного фільтра метод відеоспостереження.	Це вводить новий метод, здатний боротися з привидами, тіні і шум. Це метод краще працює в порівняння з іншими методами, а також з ним здатний полегшити проблема, пов'язана з іншим доменом.

Комплексна сегментація рухомого об’єкта на основі вейвлетів з використанням методу наближеного медіанного фільтра для відеоспостереження: представив комплексну сегментацію рухомого об’єкта на основі вейвлета з використанням методу базису приблизного медіанного фільтра. Це здатне впоратися з такими недоліками, як тінь і шум, наявні в інших методах просторової області. Ефективність цього методу оцінюється та порівнюється з іншим стандартним методом просторової області. Порівняння здійснюється за допомогою вимірювання відносної площі переднього плану,

штрафу за неправильну класифікацію, міри на основі відносного положення, нормованого хреста.

З отриманих результатів та їх якісного та кількісного аналізу можна побачити, що цей метод працює краще в порівнянні з іншими методами, а також здатний усунути проблему, пов’язану з іншими методами просторової області, такими як примари, безлад, шум тощо.

Непараметрична модель – Оцінка щільності ядра

Щоб впоратися з високочастотними варіаціями та довільними розподілами, можна використовувати непараметричні фонові моделі. Імовірність спостереження заданого значення пікселя Xt в момент t за допомогою оцінювача ядра K можна непараметрично оцінити на основі вибірки пікселів

X ={X_1, X_2, … … … X_N} as follows:

$$p(X_t) = \sum_{t=1}^{N} a_i \, K(X_t - X_i);$$

де αi — вагові коефіцієнти (зазвичай вибираються рівномірними, $\alpha i = N$ 1).

Імовірність у рівнянні можна ефективно обчислити, якщо взяти нормальну функцію N (0, ∑) як оцінювач ядра, припустивши незалежність між різними каналами кольору та використавши попередньо обчислені таблиці пошуку для функції ядра з урахуванням різниці значень інтенсивності ($Xt - Xi$) і пропускна здатність.

Використання непараметричних фонових моделей було вперше запропоновано. Щоб полегшити високі вимоги до пам’яті, пов’язані з необхідністю зберігання всього вибіркового набору кадрів, які розглядаються для оцінки щільності, введено методику оцінки, засновану на знаходженні режиму середнього зсуву. У запропоновано підхід із використанням підходу ядра змінного розміру повітряної кулі, який дозволяє уникнути оцінки параметра розміру ядра.

Однак методи оцінки щільності ядра (KDE) мають високі обчислювальні витрати. Більше того, в показано, що GMM здається кращою моделлю для простих сцен, забезпечуючи більш компактне уявлення, яке підходить для подальших кроків обробки, як, наприклад. виявлення тіні.

1.4 Сучасні тенденції та висновки

Через низьке обчислювальне навантаження віднімання фону є, мабуть, найбільш загальновизнаним початковим етапом для виявлення об’єктів ентузіазму для додатків спостереження, зокрема через використання статичних камер, і створив широке письмо. У минулих сегментах були введені первинні системи, які використовувалися для виконання цього завдання. Ці методи також були використані в багатьох інших підходах визначення, які допомагають краще впоратися з частиною труднощів, пов’язаних із підходом віднімання фону.

Ця область дає діаграму первинних шаблонів, випилюваних із фонового запису віднімання.

Очевидно, що залежно від області застосування, включаючи характеристики спостережуваних сцен і обчислювальні обмеження, найбільш підходящий підхід може відрізнятися. Дослідження різних алгоритмів віднімання фону в контексті систем моніторингу міського руху . Особлива увага приділяється компромісу між отриманими результатами та обчислювальною складністю. Підкреслено хороший компроміс, досягнутий за допомогою простих методів, таких як адаптивна медіанна фільтрація для розглянутого домену.

Представлено більш загальний вибір різних методів, що охоплюють широкий спектр основних математичних підходів. Надано категоризацію представлених підходів з урахуванням їх швидкості, вимог до пам’яті та результатів сегментації, з метою полегшення розробки/вибору підходу віднімання фону залежно від конкретних системних вимог і можливостей. Підкреслюється прийнятна точність, яка забезпечується простими методами,

такими як поточне середнє Гаусса та медіанний фільтр, висока точність моделі гауссових моделей суміші та послідовне наближення ядра за ціною вищих вимог до пам’яті та обчислень, а також проблема, пов’язана з практичною реалізація методів, що розглядають просторові кореляції.

Ініціалізація фонової моделі

Основним завданням, яке необхідно розгадати за допомогою системи віднімання фону, є встановлення моделі, яку зазвичай називають завантаженням. У контрольованих ситуаціях це час від часу досягається шляхом форсування підготовчого періоду, протягом якого помітна незаповнена сцена. У будь-якому випадку, ця методологія не підходить для загальних ситуацій спостереження. Таким чином, фонове відображення має бути введено в поле зору предметів, що рухаються. Незалежно від можливості того, що використання простих методологій, наприклад, піксельного обчислення середньої оцінки може бути достатньою для кількох застосувань, існує також незліченна кількість, зокрема, такі, що включають рої, де важливе значення має більш роз’яснений спосіб роботи зі встановленням фундаменту. Для цього зазвичай використовуються якісь просторові дані. Одна з найбільш точних методологій з огляду на це керівництво наведена, де пропонується використання даних оптичного потоку. Основна думка полягає в тому, що використання оптичного потоку в області пікселя можна уявити, щоб припустити, чи перешкоджає фоновому пікселю рухомий предмет (якщо заголовок оптичного потоку спрямований до цього пікселя) або блокований фоновий піксель розкривається (якщо оптичний потік координується далеко від цього пікселя). Стратегія, включає обробку загальної суми загальних контрастів спільно розташованого блоку зображення вхідних кадрів, щоб згрупувати їх як рухомі, статичні ближні види або статичний фон; фонове зображення складається з використанням мирського середнього каналу для з'єднання статичних фрагментів фону. Запропонована стратегія, яка передбачає ізоляцію кожного інформаційного плану у латки, які групуються

вздовж ходу подій, маючи на увазі кінцеву мету вибрати невелику кількість фонових заявників, які потім поступово вважаються фоновими чи ні. вибираючи в кожній прогресії найкраще продовження нинішньої основи, як вказують на візуальні стандарти збирання, згодом враховуючи просторові відносини, які існують всередині маленьких локалів фонового зображення. Пізніший підхід, який також розглядає з'єднання сусідніх фонових фрагментів, де консолідована рекурентна реакція квадрата-заявника та його околиці є основою вибору фрагментів, які розглядаються як фон.

Зміни освітлення та тіні

Хоча постійні зміни просвітлення ефективно забезпечуються більшістю найкращих універсальних методологій у своєму класі, раптові зміни світла та тіні, що відкидаються рухомими предметами, все ще є випробуванням для більшості з них. У зв’язку зі змінами освітлення в усьому світі, поверхневі і, тим більше, розташовані поблизу методології демонструють зміну методологій, заснованих на пікселях, завдяки яким поверхні в спостережуваній сцені досить помітні. Для прикладу відкинутих тіней усі підходи віднімання фону вказують на неадекватність, яка зазвичай виправляється на етапі після підготовки.

Раптові зміни просвітлення в усьому світі, як правило, відбуваються в просторовій обстановці. Наприклад, фреймворк, містить агентний набір фонових моделей сцени, що підходять до різних умов освітлення (незначний набір можна порівняти з увімкненим і вимкненим світлом) і вибирає модель, яка створює найменшу кількість пікселів переднього плану. Очевидно, що такий підхід вимагає попереднього вивчення вільної сцени в різних умовах освітлення. У світлі сприйняття того, що зміни освітлення можна краще врахувати з урахуванням просторових даних, структура, об’єднує результати, отримані GMM, із просторовими даними, наданими роз’єднаним просторовим поділом фону в байєсівській системі. Загальний підхід, який додатково використовує просторові зв'язки, де спостережувану сцену виправляють за допомогою методу

багатодетермінаційного перегляду світла, маючи на увазі кінцеву мету передати підготовлені відеокартини до еталонного рівня яскравості. Опціонний підхід представлений, де базова модель характеризується вимірюваною моделлю світлових впливів, а не потужністю пікселів. Крім того, ймовірність групування пікселів додатково об’єднує частини інформації поверхневого зв’язку шляхом неправильного використання поверхневих гістограм, підготовлених без зв’язку. Хоча представлені вражаючі результати, передбачається, що фон є статичним і може бути навчений заздалегідь, що є вимогою, яку можна легко виконати в сценарії, для якого розроблений підхід, доповненої реальності, але не в звичайному відеоспостереження. сценарій.

Огляд підходів до виявлення тіней представлено , де різні внески, про які повідомляється в літературі, класифікуються на чотири класи: статистичний параметричний, статистичний непараметричний, детермінований на основі моделі та детермінований не на основі моделі. З оцінюваних підходів виділяють результати. Підхід класифікує пікселі як передній план, фон, затінений фон або виділений фон, залежно від спотворення кольорів і яскравості, виміряних шляхом проектування спостережуваного значення в лінію, що проходить через початок простору RGB, і очікуване значення для кожного пікселя. позицію. Підхід класифікує пікселі як передній план або фон в залежності від відстані в колірному просторі HSV спостережуваного до очікуваних значень для кожної позиції пікселя, таким чином використовується різний вплив, який умови освітлення мають на відтінок, насиченість і значення каналів.

2 АЛГОРИТМІЧНІ ТА МАТЕМАТИЧНІ РІШЕННЯ ДЛЯ ЗАДАЧ ДАНОГО КЛАСУ

Цей розділ служить для окреслення роботи, виконаної під час дипломної роботи, і опису того, як були отримані результати. Першим кроком роботи було ознайомлення з комп'ютерним баченням як галуззю дослідження та найсучаснішим відстеженням руху. Це було зроблено шляхом огляду літератури, особливо було використано для пошуку відповідних кандидатів для оцінки. З цих двох найсучасніших трекерів було обрано модель гаусової суміші (GMM) та оцінку щільності ядра (KDE) . На додаток до цього, для створення відеоінженерії, за допомогою алгоритму віднімання фону ми віднімаємо фон і передній план з різних відео, а потім можемо змінити або фон, або об'єкт переднього плану на новий.

Розроблено систему для проведення тестування алгоритмів, написану на C++ з використанням бібліотеки OpenCV для обробки зображень. Реалізації GMM та KDE є незначними модифікаціями загальнодоступного коду. Код був змінений, щоб забезпечити узгоджений інтерфейс для всіх алгоритмів відстеження та зробити можливим їх використання разом із відніманням фону. Результати конкурсу доступні, тому можна порівняти результати реалізацій з цієї дипломної роботи з результатами авторів оригінального алгоритму.

Одна реалізація оригінального алгоритму доступна як частина BGSLibrary3, реалізація, використана при тестуванні в цій дипломній роботі, повністю заснована на написаній статті та модифікована для роботи з рухомою камерою. Не всі функції, описані були реалізовані.

Два трекери GMM та KDE були оцінені, щоб визначити, чи виграють вони від фонового віднімання. Детальна інформація про оцінку наведена в наступному розділі.

2.1 Віднімання фону

Отримана маска фону від фонового віднімання використовується різними способами для трекерів. Для GMM і KDE нове зображення створюється з вихідного зображення шляхом встановлення чорних фонових пікселів. Ефективність фонового віднімання оцінюється в різних випадках, щоб побачити, що впливає на його продуктивність. Побудовано два випадки.

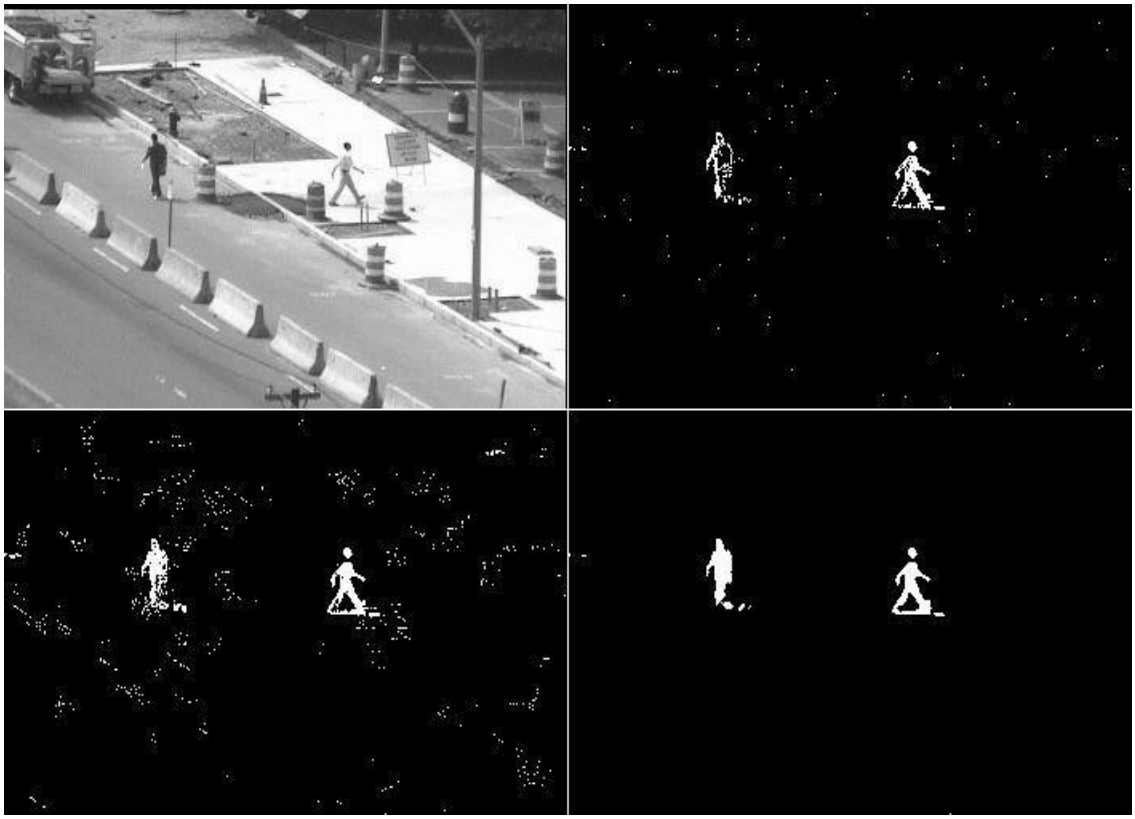


Рисунок 2.1 - Приклад віднімання фону

По-перше, це проста послідовність, яка складається з одного великого зображення та створення відео, прогортаючи його за допомогою меншого вікна. У цьому випадку всі пікселі слід класифікувати як фон, оскільки в статичному зображенні немає рухомих об’єктів.

Другий випадок — оцінити результат віднімання, коли в даних для руху камери немає помилки. Це було зроблено шляхом запису послідовності без переміщення камери, а потім створення нового відео з використанням невеликих

частин оригінальної послідовності (переміщення вікна над ним, імітація рухомої камери). Роблячи це, ми мінімізуємо вібрації та отримуємо досконалі знання про рух за кадр.

Нарешті, віднімання фону оцінюється на послідовностях з камери за умов: лише рухи панорамування, лише рухи нахилу та рухи як панорамування, так і нахилу. Для цієї фонової моделі припускаються невеликі кутові повороти, а рух камери апроксимується як трансляція.

Параметри камери

Щоб виконати віднімання фону, коли джерело кадрів рухається, потрібно дізнатися про піксельний баланс між кадрами. Один з підходів, щоб отримати це без попереднього вивчення розробки, полягає в тому, щоб слідувати фокусам, які мають місце, починаючи з одного корпусу, а потім на наступний, використовуючи оптичний потік.

У той момент, коли стає доступним вивчення руху камери, потрібен певний підхід, щоб зв'язати зміни контейнера та кута нахилу зі змінами положення пікселя на зображенні. У той момент, коли коригування в контейнері та нахилі невеликі, коригування положення пікселя можна приблизно оцінити відносно зміни положення.

$$\Delta x = c1 \,.\, \Delta p$$
$$\Delta y = c2 \,.\, \Delta t$$

Коефіцієнти $c1$ і $c2$ можна оцінити, наприклад, за допомогою оптичного потоку, щоб отримати оцінку руху пікселя і порівняти це зі зміною панорамування та нахилу. Їх також можна знайти, вручну зіставивши зображення з відомим положенням камери та обчисливши з цього коефіцієнти. Обидва методи оцінюються.

Тестові процедури для оцінки трекерів із фоновим відніманням

Тестовий приклад 1: відстеження з фоновим відніманням.

Розташування кадрів збирається зі статичної камери. У серії є не менше одного рухомого об’єкта. Камера, що рухається, представлена створенням іншої групи кадрів, де кожен кадр є площею встановленого розміру від відповідного статичного контуру камери. Позиція пікселя вилученої локалі реєструється для імітації даних інтерпретації від робота. Потім для кожного трекера:

• Ініціювати обмежувальну рамку на об'єкті, який слід відстежувати.

• Відстеження з відніманням фону та без нього.

• Якщо відстеження втрачено, повторіть ініціалізацію, надавши нову обмежувальну рамку навколо об’єкта.

• Порахуйте, скільки разів втрачено відстеження та скільки кадрів обробляється за секунду.

Тестовий приклад 2: просте відстеження з відніманням фону та роботом. Послідовності кадрів збираються з камери, в послідовності перед простим фоном є рухоме одноколірне коло. Позиції панорамування та нахилу реєструються, щоб можна було вирівняти рамки. Послідовності:

• Камера панорамується лише вправо та вліво.

• Камера нахилена лише вгору та вниз.

• Камера панорамується та нахиляється нерегулярно.

• Об'єкт, що рухається, переміщується нерегулярно, камерою керують вручну, щоб утримувати об'єкт у центрі зображення (як це керує трекер).

Для кожного трекера:

• Ініціювати обмежувальну рамку на об'єкті, який слід відстежувати.

• Відстеження з відніманням фону та без нього.

• Якщо відстеження втрачено, повторіть ініціалізацію, надавши нову обмежувальну рамку навколо об’єкта.

• Порахуйте, скільки разів втрачено відстеження та скільки кадрів обробляється за секунду.

Тестовий приклад 3: відстеження за допомогою віднімання фону та роботи

• Така ж процедура, як сценарій 2, але з більш складним фоном із безладом.

2.2 Алгоритм моделі гаусової суміші

Щоб врахувати складні фони, що містять більше одного гауссового розподілу, [28] моделює кожен піксель як суміш K гауссів, що відповідає будь-якому фону

$$P(I_{p,t}) = \sum_{i=1}^{K} w_{i,p,t} * \eta(I_{p,t}\ ;\ \mu_{i,p,t})$$

де η($\mu i,p,t$, $\sum i,p,t$) — фонова модель Гаусса ith, а $\omega i,p,t$ її вага. Значення пікселів, які не відповідають розподілу фону, розглядаються як передній план, доки не буде достатніх і послідовних доказів, щоб ініціювати новий гаусів для їх підтримки. Фонові гауссіани можна визначити з точки зору його стійкості та дисперсії, яку можна виміряти за допомогою ω/σ. Це значення збільшується в міру того, як розподіл стає більш впевненим і стійким. Після впорядкування гауссів за ω/σ, перші B-розподіли вибираються як фонова модель

$$B_{p,t} = \arg\min_{b}(\sum_{i=1}^{b} \omega_{i,p,t} > T)$$

де T є мірою мінімальної частини даних, яка має належати до фону. Таким чином, Ip,t позначається як фон, якщо це стандартне відхилення фонової моделі Гаусса. ГММ набув величезної популярності. Проте не вдається досягти чутливого виявлення у випадку, коли фон має дуже високі частотні варіації, такі як хвилювання води або струшування листя дерев, тобто фон, який має швидкі варіації, не може бути точно змодельований лише за допомогою кількох гауссів. Ще одним важливим моментом є його здатність адаптуватися до раптової зміни фону, яка залежить від швидкості навчання. Низька швидкість навчання підходить для довгострокових змін, але вона погано адаптується до раптових змін. Висока швидкість навчання може швидко адаптуватися до змін.



Рисунок 2.2 - Приклад моделі гаусової суміші

2.3 Алгоритм оцінки щільності ядра

Метод оцінки щільності ядра (KDE), непараметричний підхід, який може ефективно адаптуватися до динамічного фону. У кожному пікселі KDE обчислюється за таким рівнянням під час індексу часу t:

$$p(x) = \frac{1}{n}\sum_{i=1}^{n} K(x - x_t)$$

де n – кількість загальної кількості спостережуваних кадрів, а xt – спостережуване значення під час індексу часу t. p(x) — середнє значення нормальних густин з центром у зразку x. Функція ядра K(x) повинна задовольняти наступним умовам: використовується як функція ядра. У дослідженні, проведеному Парком та ін., було зібрано багато кадрів, перш ніж оцінити модель фону Гаусса, і, таким чином, потрібен був великий обсяг пам’яті. Щоб подолати цей недолік, ми модифікуємо оригінальний метод KDE і пропонуємо схему, яка використовує перший кадр для ініціалізації фонової моделі KDE. У першому кадрі більшість пікселів представляють фон, а в деяких

інших пікселях є передній план. Навіть якщо ми використали перший кадр для ініціалізації фонової моделі, інформація переднього плану буде зменшена і залишиться лише фоновою інформацією через процес оновлення, оскільки фонові значення зустрічаються частіше, ніж значення переднього плану на рівні пікселя. Гауссова модель KDE згодом оновлюється на кожному кадрі, керуючи швидкістю навчання відповідно до ситуації. Імовірність $pt(x)$ базується на кожному пікселі і може бути виражена як:

$$p_t(x) = \hat{p}_{t-1}(x) + \frac{1}{G_t\sqrt{2\pi\sigma^2}} \exp\left[\frac{1}{2}\left(\frac{x - x_t}{\sigma}\right)^2\right]$$

Кожен піксель має імовірнісну модель. Імовірність, отримана методом KDE, додається до попередньої щільності ймовірності в кожному кадрі. У другому рівнянні Gt використовується як швидкість навчання в момент t і може бути змінена залежно від таких факторів, як зміна часу та освітлення. Оскільки ймовірність повинна задовольняти $\int pt(x)dx = 1$, $pt(x)$ нормується таким чином:

$$\hat{p}_t(x) = p_t(x) \,/\, \sum_{x=0}^{N} p_t\,(x)$$

де $pt(x)$ — нормальна щільність у зразку x та в часовому індексі t. $\hat{pt}(x)$ — нормована нормальна щільність, а N — загальна кількість зразків. За допомогою вищезгаданого процесу отримують нову імовірнісну фонову модель. Цей метод оновлення покращує ефективність пам’яті, оскільки для ініціалізації ймовірнісної фонової моделі не потрібно зберігати багато зображень. Метод оновлення автоматично зменшує ймовірність появи неважливих фонів, які не з’являються протягом тривалого періоду, шляхом додавання додаткової ймовірності та виконання кроку нормалізації. Наприклад, коли автомобіль, припаркований на тривалий час, рухається або зникає, запропонований метод

постійно оновлює середовище. Таким чином, з'являється нова довідкова інформація, а попередня неважлива фонова ймовірність, пов'язана з автомобілем, автоматично знижується шляхом оновлення фонової моделі. Ми використовували Gt як параметр для контролю швидкості навчання. Якщо Gt збільшується, нова інформація повільно вивчається, а попередня інформація повільно зникає. Якщо Gt зменшено, алгоритм швидко адаптується до середовища та швидко видаляє стару інформацію. На початковому етапі фонова модель повинна швидко адаптуватися до нового середовища і, з часом, фон повинен мати стабільний процес оновлення. З цієї причини Gt було використано як сигмовидна функція, яка може бути виражена наступним чином:

$$G_t = Gain * \frac{2}{1 + \exp(-(cnt - \beta)/\lambda}$$

Деякі проблеми, пов'язані з непараметричним підходом до оцінки щільності ядра, - це небажано довгий час обробки та велика потреба в пам'яті. Ми можемо зменшити складність і потребу в пам'яті, використовуючи наближення гістограми. Гауссова ймовірність і приклад наближення гістограми. Bd – це ширина гістограм за розміром d, Ck – центр кожної гістограми, а k – номер гістограми? Параметр Bd можна розрахувати за таким рівнянням :

$$B_d = \frac{\max(x^d) - \min(x^d)}{N_d} \qquad d = 1,2,3$$

де Nd представляє кількість бінів для кожного виміру d, а xd — значення пікселя у вимірі d.



Рисунок 2.3 - Приклад KDE

Виявлення тіней

Для видалення тіней від рухомих об’єктів ми застосували алгоритм визначення рухомої відкинутої тіні, який виявився досить точним і придатним для усунення тіней. Основна ідея полягає в тому, що відкидана тінь затемнює фон, при цьому колір самого фону не змінюється. Використовуючи цей принцип, ми можемо виразити алгоритм видалення тіні так

$$[\rho \leq \frac{x^v}{Bg^v} \leq \delta] \wedge (|x^s - Bg^s| \leq \tau_s) \wedge (|x^h - Bg^h| \leq \tau_h)$$

де Bgh, Bgs та Bgv представляють відтінок, насиченість та компоненти освітлення відповідно фонових пікселів із значеннями фону, які є найближчими до вхідного зображення серед моделей фонової гістограми. xv, xs і xh представляють компоненти відтінку, насиченості та освітлення вхідних відеопікселів.

Адаптація до раптової зміни освітлення

Якщо сам фон значно змінений (наприклад, раптово посвітлішав або затемнів), потрібна швидка адаптація. Ми можемо отримати цей ефект,

ініціалізувавши значення cnt. Якщо значення cnt ініціалізується, Gt також ініціалізується і швидкість адаптації до фону збільшується:

$$\begin{Bmatrix} Mv_t = (G_t - 1 * \frac{Mv_{t-1}}{G_t} + mean_{\forall_{i,j}}(Dist_v(i,j))/G_t \\ if\left(\left|mean_{\forall_{i,j}}(Dist_v(i,j)) - Mv_t\right| > T_v\right)\ then, cnt = {}^{\beta}/_{2} \end{Bmatrix}$$

Tv є порогом для ініціалізації cnt; в наших експериментах встановлено значення 30. Distv(i, j) — це значення освітлення поточного вхідного зображення на (i, j) пікселі. Mvt – це ковзне середнє значення середньогоꓯi,j(Distv(i, j)).S.

Традиційне відеоспостереження сьогодні широко застосовується на об'єктах транспортної інфраструктури. Як правило, сучасні оглядові відеокамери добре справляються з роботою в складних умовах освітленості, оскільки мають високу оптичну здатність (HD) і використовуються спільно з датчиками руху. З їх допомогою співробітники служби безпеки відстежують обстановку і самостійно приймають рішення про те, як діяти при виникненні нештатної ситуації. За допомогою відео-камер здійснюється як цілодобовий моніторинг в реальному часі, так і ретроспективний пошук по архіву даних. Але в роботі традиційного відеоспостереження є величезний мінус – занадто багато залежить від людського фактора. Оператори змушені безперервно відстежувати весь відеопотік, ризикуючи пропустити щось важливе, особливо у кінці зміни, коли уважність персоналу знижується. Додамо до цього і складність пошуку інформації в «сирому», ще не автоматизованому архіві: іноді необхідно оглядати години відео, навіть якщо приблизний час події відомий. Тому на великих об'єктах транспортної інфраструктури – пересадочних вузлах, вокзалах, аеропортах тощо – вже практично повсюдно застосовується не звичайне відеоспостереження, а комплексні інтелектуальні системи, в основі яких лежать складні алгоритми відеоаналітики. Такі завдання, як фіксація

руху та автоматичне виявлення залишених предметів, належать до базової функціональності подібних технологій, але можливості їх набагато ширше. Відчутне зростання кількості впроваджень таких систем, навіть в умовах економічної кризи, на різних інфраструктурних, розважальних, спортивних та інших об'єктах показує високу потребу в розумних технологіях відеоспостереження для забезпечення безпеки.

Відеоаналітика на транспорті застосовується на декількох рівнях. До базових відносять завдання розпізнавання стандартизованих образів, наприклад номерів транспортних засобів. Слід зазначити, що для експлуатації в реальних умовах потрібно не тільки точність розпізнавання номерних знаків, а й надійність у важких погодних умовах, захищеність та інші характеристики [1].

Більш складним завданням є розпізнавання осіб, коли від системи потрібно, по-перше, виділити особу в натовпі, а потім однозначно визначити (ідентифікувати) конкретну людину, порівнявши зображення, отримане з камери, з фотографією в базі даних.Критеріями якості роботи таких систем є точність (частка правильно ідентифікованих і пропущених осіб), швидкість розпізнавання, а також час на пошук і порівняння з особами з бази даних. Ефективні системи повинні показувати хороші результати навіть в умовах поганого освітлення, при несприятливих погодних умовах і в щільному потоці людей. Використання нейронних мереж і алгоритмів глибокого навчання дуже добре справляється з такими завданнями.

Ще вище рівнем стоять завдання, які формулюються вже не так однозначно. Наприклад досить часто необхідно виявити неадекватну поведінку людини. Однозначно визначити параметри такої оцінки вкрай важко, адже «адекватність» поведінкових установок в різних обставинах може трактуватися по-різному. Розмитість критеріїв веде до нечітких правил прийняття рішень і, як наслідок, – до неефективних реакцій. Створення систем такого класу є справою майбутнього, проте в цьому напрямку багато вже зроблено. Наприклад, системи відеоспостереження на об'єктах транспортної інфраструктури здатні

відстежувати рухомих людей, що розглядаються як потенційно більш небезпечні об'єкти. На практиці таке завдання вимагає складного технічного рішення. Наприклад, як відстежити переміщення конкретної людини у величезному переповненому залі очікування, де одночасно з ним в різних напрямках рухається безліч людей? Зазначимо, що стандартне відеоспостереження повною мірою з таким завданням не впорається, в той час як сучасні алгоритми відеоаналізу цю роботу здатні виконати.

Впровадження технологій інтелектуального аналізу має бути невід'ємною комплексною частиною модернізації всієї транспортної інфраструктури. Саме так зараз відбувається в багатьох країнах світу, де не тільки комунікації, системи управління рухом та сигналізацією, рухомий склад і транспортна мережа є об'єктами модернізації, але також йде активне впровадження більш досконалих інструментів об'єктивного контролю і реагування. І це особливо важливо в умовах зростаючих загроз безпеки: необхідно не постфактум ліквідувати наслідки, а запобігати самій загрозі виникнення нештатних ситуацій, які можуть бути від зловмисних злочинних дій або є проявом недбалості персоналу. Всі сучасні транспортні термінали проектуються з урахуванням обладнання приладами відеоспостереження, однак обладнання тих об'єктів транспортної інфракструктури, які були побудовані за часів, коли відеоспостереження ще не застосовувалося, пов'язане з певними труднощами. Прикладом є звичайні пішохідні переходи. Це оптимальне місце установки інтелектуальних камер розпізнавання осіб, оскільки особу зручніше розпізнати саме тоді, коли людина рухається в направленому потоці по певній траєкторії. Але тут виникає серйозна проблема: як забезпечити достатню освітленість. Якщо для оглядового відеоспостереження рівень освітленості в 100 лк достатній, то для штатної роботи систем дистанційного біометричного розпізнавання осіб зі стандартними камерами світла потрібно істотно більше. Оскільки умови на кожному об'єкті відрізняються, проектування таких систем має індивідуальний характер, як правило, технічні рішення нестандартні і поодинокі. Одним з виходів може стати

використання спеціалізованих камер відеоспостереження, менш вимогливих до умов освітлення. Сучасні аналітичні системи відеоспостереження для транспорту будуються з урахуванням охоплення значної території, на якій необхідно забезпечити цілодобовий моніторинг і мінімізувати ризики проникнення всередину охоронюваних периметрів і прилеглої території. Оскільки транспортні вузли мають складну розгалужену структуру, застосування спеціальних технічних засобів є вкрай важливим інструментом скорочення часу реагування на нештатні ситуації.

3 ВЛАСНЕ ОРИГІНАЛЬНЕ АЛГОРИТМІЧНЕ ТА ПРОГРАМНЕ РІШЕННЯ ЗАДАЧІ

У цьому розділі ми проведемо ряд експериментів. Ми використовуємо GMM і KDE для моделювання фону, для віднімання і для відеоінженерії, для оновлення фону або об’єкта переднього плану. Проте, як ми обговорювали в розділах 2 за допомогою найсучасніших і детальних алгоритмів.

Для аналізу ефективності підходів віднімання фону в різних середовищах використовувалися різні конкуруючі відеопослідовності з роздільною здатністю 240 × 320 при 30 кадрах в секунду.

Результати виявлення представлені якісно та кількісно. Параметри для кожного алгоритму були визначені експериментально. Для кожної послідовності представлені кілька репрезентативних кадрів, основна істина та результати виявлення, отримані кожним алгоритмом. Результати виявлення відображаються у вигляді чорно-білих зображень, де білі пікселі представляють об'єкти переднього плану, а чорні пікселі представляють фон. Ефективність кожного підходу також оцінюється кількісно за допомогою а) традиційних метрик оцінки на основі пікселів (точність, відкликання, F-виміри), які зазвичай використовуються при оцінці підходів віднімання фону, і b) оціночних метрик на основі компонентів, розроблених з з точки зору виявлення об'єктів, тут ми використовуємо правильну швидкість виявлення, частоту виявлення промахів і частоту помилкових.

3.1 Показники оцінки

У цій дипломній роботі ми використовуємо два типи вимірювань для оцінки ефективності різних підходів, один визначений на рівні пікселів, інший — на рівні компонентів.

Перший тип оціночних показників, визначених на рівні пікселя, є найбільш прямим виміром, який часто використовується для оцінки ефективності підходів віднімання фону, включаючи точність, відкликання та F-міру. Вони визначаються таким чином:

$$Precision = \frac{\#True\ Positives}{\#true\ Positives + \#False\ Postives}$$

$$Recall = \frac{\#True\ Postives}{\#True\ Positives + \#False\ Negatives}$$

$$F - measure = 2\,.\frac{Precision\,.recall}{Precision + Recall}$$

Ці показники оцінки вимірюють точність підходу на рівні пікселя, однак у деяких випадках людей цікавить не виявлення точкових цілей, а регіони об’єктів.

Таким чином, ми також використовуємо об’єктно-орієнтовані метрики оцінки.

Якщо бути більш конкретним, ми розглянемо три випадки, які показані таким чином:

• Правильне виявлення (CD) або збіг 1-1: виявлена область відповідає одній і тільки одній основній області істинності.

• Помилкова тривога (FA): виявлена область не має відповідності в основній істині.

• Помилка виявлення (DF): основна область істини не виявлена.

Відповідно до визначень, нам потрібно визначити відповідність області переднього плану в результаті виявлення та в основній істині, тобто чи відповідає область переднього плану в основній істині з сегментацією. На основі

відповідності ми можемо оцінити обраний підхід з точки зору правильної частоти виявлення, частоти помилкових тривог та частоти помилок виявлення.

3.2 Результати експерименту

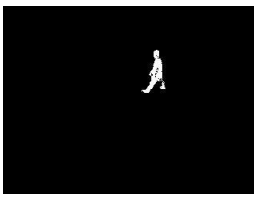
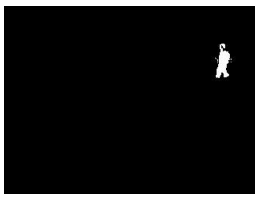
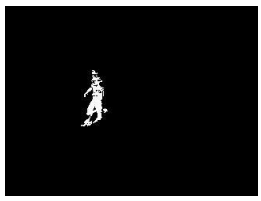
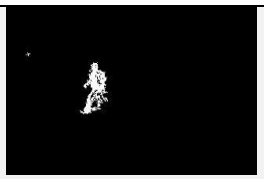
У цьому розділі представлені експериментальні результати, що порівнюють різні підходи алгоритмів віднімання фону. Експерименти проводяться на різних послідовностях, які демонструють, що наш підхід випереджає ці алгоритми, і він стійкий до відхилень від неточних оцінок руху та зміщення пікселів під час реєстрації послідовних зображень.

Результат порівняння підходів, заснованих на зовнішньому вигляді, з підходами, що включають рух і зовнішній вигляд, демонструє, що рух може забезпечити більшу розрізнювальну силу, ніж використання окремої підказки зовнішнього вигляду, що може підвищити стійкість до викидів від реєстрації зображення, але разом з моделюванням руху і сигналів зовнішнього вигляду є вразливим до цих викидів з будь-якого сигналу, оскільки ці викиди можуть бути введені в функцію спільного ядра, що погіршить її точність.

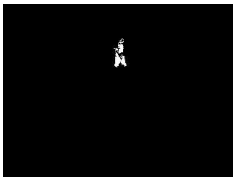
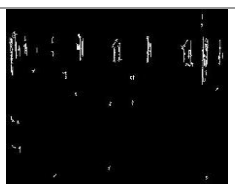
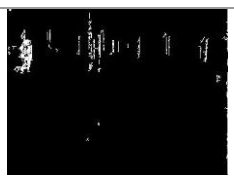
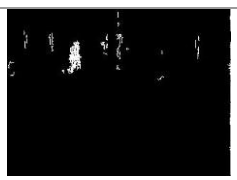
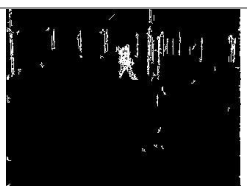
Оцінка граничних ймовірностей корисна для боротьби з викидами та забезпечує більшу точність, але відкликання може бути набагато нижчим, оскільки воно може бути надмірно консервативним.

Фонові результати віднімання

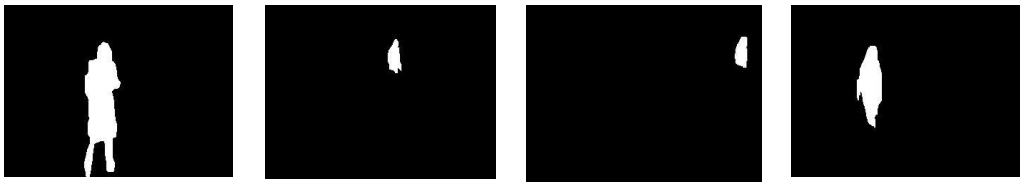
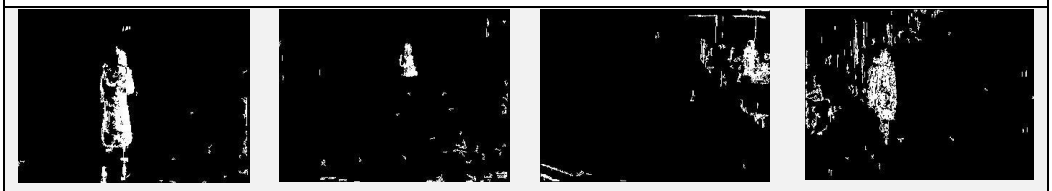
Таблиця 3.1 - Якісне порівняння відео послідовності 1

ith	$155st$	$221st$	$591st$	$811st$
IMG				
GMM				
KDE				

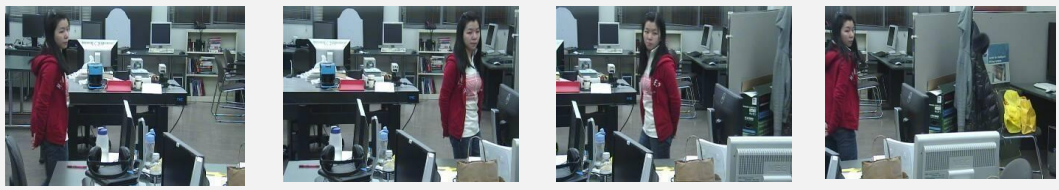
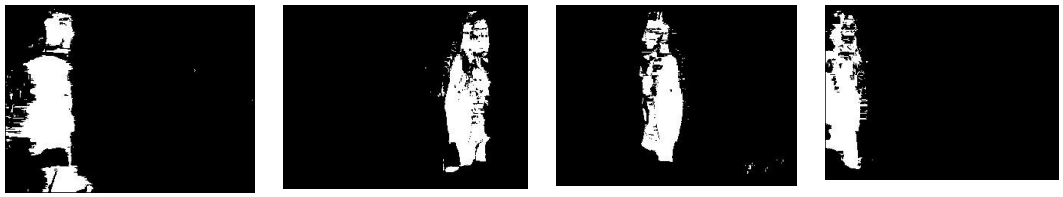
Таблиця 3.2 - Якісне порівняння відеоряду 2

ith	125th	237th	272nd	305th
IMG				
GMM				
KDE				

Таблиця 3.3 - Якісне порівняння послідовності в приміщенні 3.

IMG	
GMM KDE	 

Таблиця 3.4 - Якісне порівняння послідовності 4 в приміщенні

IMG	
GMM KDE	 



(a) (b) (c)

Рисунок 3.1 (a) Містить об'єкт переднього плану як вхід (b), що представляє нерухомість або фон сцени (c), що представляє вихід з новим об'єктом переднього плану.



Рисунок 3.2 - Результат фонового оновлення

3.3 Обговорення

У цьому розділі представлені експериментальні результати, що порівнюють два підходи до вилучення фону. Експерименти проводяться на

різних відеопослідовностях, які демонструють, що підхід GMM перевершує інший алгоритм і стійкий до викидів, що виникають через неточну оцінку руху та зміщення пікселів при реєстрації послідовних зображень. Результат порівняння підходів, заснованих на зовнішньому вигляді, з підходами, що включають рух і зовнішній вигляд, демонструє, що рух може забезпечити більш високу розрізнювальну силу, ніж використання окремої підказки зовнішнього вигляду, що може покращити стійкість до викидів від реєстрації зображення, але моделює рух і сигнали появи. jointly вразливий до цих викидів з будь-якого сигналу, оскільки ці викиди можуть бути введені в функцію спільного ядра, що погіршить її точність.

Крім того, також показані результати відеоінженерії, в якій ми робили монтаж відео, наприклад, у нас є два відео, за допомогою яких ми змінили фон одного відео на інше, а в одному відео ми змінили об’єкти переднього плану на нові об’єкти з таким же фоном. Відеоінженерія в основному здійснювалася ручним методом, але в нашому випадку ми робили це автоматично. Ми передали відео в алгоритм, і він автоматично замінює об’єкти. Цей метод може бути корисним для кіноіндустрії, а також у сфері відеоспостереження.

ВИСНОВКИ

В дипломній роботі проводиться дослідження та виявлення об’єктів довільного візуального вигляду у відео даних спостереження. Зокрема, об’єкти, що цікавлять, були двох різних характерів: рухомі об’єкти, які проходять повз спостережувану сцену, та статичні об’єкти, які додаються або видаляються зі сцени.

Об'єкти, що рухаються, повинні бути надані на рівні аналізу вищого рівня для розпізнавання дій і програмної поведінки. Статичні об'єкти повинні надавати онлайн-попередження операторам в режимі реального часу. Відсутність моделей зовнішнього вигляду (і неможливість їх побудови) і нерухомість статичних об’єктів призвели до використання віднімання фону як інструмента низькорівневої обробки. Надано ретельний огляд найсучасніших методів віднімання фону, таким чином висвітлюючи основні проблеми, з якими стикається ця техніка, і те, як ці проблеми розглядаються у літературі.

В якості базових фонових моделей було вибрано моделі гаусової суміші і модель щільності ядра. У цій роботі віднімання фону було виконано обома методами, і ми виявили, що перший є кращим підходом для віднімання фону, і з його допомогою ми також змогли виконати деяку відеоінженерію. Ми замінили об’єкт переднього плану відео на інший об’єкт, витягнутий з іншого відео.

Відеоінженерія зазвичай ручний процес, але за допомогою цього методу можна зробити це автоматично. Єдина ручна частина нашого підходу — це надати вхідні відео для методу. Відеоінженерія також може бути корисною для редагування відео шляхом внесення змін у відео.

У пропонованому підході для параметричної моделі виконується віднімання фонових моделей і показуються їх результати. Другий підхід стосується непараметричних моделей, де фонове віднімання виконується за допомогою непараметричної моделі.

Крім того, ми зробили деяку відеоінженерію, де використовувався алгоритм віднімання фону, щоб фон з одного відео та переднього плану з іншого

відео були об’єднані, щоб утворити нове відео. Роблячи таким чином, ми також можемо робити більш складну відеоінженерію з кількома відео. Нарешті, результати аналізу регіонів можуть бути використані для покращення якості фонових моделей, таким чином, значно покращуючи результати.

СПИСОК ПОСИЛАНЬ НА ДЖЕРЕЛА

[1] C. Anderson, Peter Burt, and G. van der Wal. Change detection and tracking using pyramid

transformation techniques. In Proceedings of SPIE - Intelligent Robots and Computer Vision,

volume 579, pages 72–78, 1985.

[2] American Society of Photogrammetry ASP. Manual of Photogrammetry. Fourth Edition, American Society of Photogrammetry, Falls Church, 1980.

[3] J. Barron, D. Fleet, and S. Beauchemin. Performance of optical flow techniques. International Journal of Computer Vision, 12(1):42–77, 1994.

[4] J. Bergen, P. Anandan, K. Hanna, and R. Hingorani. Hierarchical model-based motion estimation. In Proceedings of the European Conference on Computer Vision, 1992.

[5] K. Bradshaw, I. Reid, and D. Murray. The active recovery of 3d motion trajectories and their use in prediction. IEEE Transactions on Pattern Analysis and Machine Intelligence, 19(3):219–234, March 1997.

[6] R. Collins, Y. Tsin, J.R. Miller, and A. Lipton. Using a DEM to determine geospatial object trajectories. In Proceedings of the 1998 DARPA Image Understanding Workshop, pages 115– 122, November 1998.

[7] R.T. Collins and Y. Tsin. Calibration of an outdoor active camera system. In Proceedings of the 1999 Conference on Computer Vision and Pattern Recognition, pages 528–534. IEEE Computer Society, June 1999.

[8] F. Dellaert and R.T. Collins. Fast image-based tracking by selective pixel integration. In ICCV99 Workshop on Frame-Rate Applications, September 1999.

[9] B. Flinchbaugh and T. Bannon. Autonomous scene monitoring system. In Proc. 10th Annual Joint Government-Industry Security Technology Symposium. American Defense Preparedness Association, June 1994.

[10] H. Fujiyoshi and A. Lipton. Real-time human motion analysis by image skeletonization. In Proceedings of the 1998 Workshop on Applications of Computer Vision, 1998. Robotics Institute, CMU – 66 – VSAM Final Report

[11] Geometric Geodesy Branch GGB. Geodetic Survey. Publication SMWD3-96-022, Phase II, Interim Terrain Data, Fort Benning, Georgia, May 1996.

[12] M. Hansen, P. Anandan, K. Dana, G. van der Wal, and P. Burt. Real-time scene stabilization

and mosaic construction. In Proc. Workshop on Applications of Computer Vision, 1994.

[13] I. Haritaoglu, Larry S. Davis, and D. Harwood. W4 who? when? where? what? a real time system for detecing and tracking people. In FGR98, 1998.

[14] M. Isard and A. Blake. Contour tracking by stochastic propagation of conditional density. In Proceedings of the 1996 European Conference on Computer Vision, pages 343–356, 1996.

[15] Institute for Simulation & Training IST. Standard for Distributed Interactive Simulation – Application Protocols, Version 2.0. University of Central Florida, Division of Sponsored Research, March 1994.

[16] S. Ju, M. Black, and Y. Yacoob. Cardboard people: A parameterized model of articulated image motion. In Proceedings of International Conference on Face and Gesture Analysis, 1996.

[17] T. Kanade, R. Collins, A. Lipton, P. Anandan, and P. Burt. Cooperative multisensor video surveillance. In Proceedings of the 1997 DARPA Image Understanding Workshop, volume 1, pages 3–10, May 1997.

[18] T. Kanade, R. Collins, A. Lipton, P. Burt, and L. Wixson. Advances in cooperative multisensor video surveillance. In Proceedings of the 1998 DARPA Image Understanding Workshop, volume 1, pages 3–24, November 1998.

[19] D. Koller, K. Daniilidis, and H. Nagel. Model-based object tracking in monocular image sequences of road traffic scenes. International Journal of Computer Vision, 10(3):257–281, June 1993.

[20] A. Lipton, H. Fujiyoshi, and R.S. Patil. Moving target detection and classification from realtime video. In Proceedings of the 1998 Workshop on Applications of Computer Vision, 1998.

[21] Alan J. Lipton. Local application of optic flow to analyse rigid versus non-rigid motion. In ICCV99 Workshop on Frame-Rate Applications, September 1999.

[22] T. Matsuyama. Cooperative distributed vision. In Proceedings of DARPA Image Understanding Workshop, volume 1, pages 365–384, November 1998.

[23] H. S. Sawhney, S. Hsu, and R. Kumar. Robust video mosaicing through topology inference and local to global alignment. In Proc. European Conference on Computer Vision, 1998.

[24] H. S. Sawhney and R. Kumar. True multi-image alignment and its application to mosaicing and lens distortion. In Proc. IEEE Conference on Computer Vision and Pattern Recognition, 1997.

Robotics Institute, CMU – 67 – VSAM Final Report

[25] C. Tomasi and T. Kanade. Shape and motion from image streams: factorization method. International Journal of Computer Vision, 9(2), 1992.

[26] K. Toyama, J. Krumm, B. Brumitt, and B. Meyers. Wallflower: Principles and practice of background maintenance. In Proc. International Conference on Computer Vision, pages 255–261, 1999.

[27] L. Wixson, J. Eledath, M. Hansen, R. Mandelbaum, and D. Mishra. Image alignment for precise camera fixation and aim. In Proc. IEEE Conference on Computer Vision and Pattern Recognition, 1998.

[28] L. Wixson and A. Selinger. Classifying moving objects as rigid or non-rigid. In Proc. DARPA Image Understanding Workshop, 1998.

[29] C. Wren, A. Azarbayejani, T. Darrell, and Alex Pentland. Pfinder: Real-time tracking of the human body. IEEE Transactions on Pattern Analysis and Machine Intelligence, 19(7):780– 785, 1997

БІБЛІОГРАФІЧНА ДОВІДКА

Тема дипломної роботи бакалавра: «Інформаційні моделі та засоби роботи систем відео моніторингу»

Обсяг пояснювальної записки: 61 аркуш

Дата закінчення дипломної роботи 13 червня 2022р.

Підпис студента _____