

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
СХІДНОУКРАЇНСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМ. В. ДАЛЯ
ФАКУЛЬТЕТ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ ТА ЕЛЕКТРОНІКИ
КАФЕДРА КОМП'ЮТЕРНИХ НАУК ТА ІНЖЕНЕРІЇ

До захисту допускається
Завідувач кафедри
_____ Скарга-Бандурова І.С.
«_____» _____ 20__ р.

МАГІСТЕРСЬКА РОБОТА

НА ТЕМУ:

**Методи та інформаційна технологія виявлення та класифікації двомірних
об'єктів**

Освітній ступінь “Магістр”
Спеціальність 122 “Комп’ютерні науки”

Науковий керівник роботи:

(підпис)

Д.О.Недзельський

(ініціали, прізвище)

Консультант з охорони праці:

(підпис)

Я.О.Критська

(ініціали, прізвище)

Студент:

(підпис)

О.Є. Березін

(ініціали, прізвище)

Група:

КН-17зм

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
СХІДНОУКРАЇНСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ІМЕНІ ВОЛОДИМИРА ДАЛЯ

Факультет Інформаційних технологій та електроніки

Кафедра Комп'ютерних наук та інженерії

Освітній ступінь магістр

Напрямок підготовки _____

(шифр і назва)

Спеціальність 122 "Комп'ютерні науки"

(шифр і назва)

ЗАТВЕРДЖУЮ:

Завідувач кафедри _____

І.С. Скарга-Бандурова

« _____ » _____ 20 ____ р.

**З А В Д А Н Н Я
НА МАГІСТЕРСЬКУ РОБОТУ СТУДЕНТУ**

Березіну Олександрю Євгеновичу

(прізвище, ім'я, по батькові)

1. Тема роботи Методи та інформаційна технологія виявлення та класифікації двомірних об'єктів

керівник проекту (роботи) Недзельський Дмитро Олександрович, к.т.н., доц.

(прізвище, м. 'я, по батькові, науковий ступінь, вчене звання)

затверджені наказом вищого навчального закладу від «18» 10 2018 р. № 221/48

2. Строк подання студентом роботи 10.01.2018

3. Вихідні дані до роботи Матеріали науково-дослідної практики, тестові зображення растрових форматів, методи пошуку двомірних об'єктів, теоретичні відомості про методи пошуку двомірних об'єктів

4. Зміст розрахунково-пояснювальної записки (перелік питань, які потрібно розробити) Аналіз предметної області, розробка алгоритму інформаційної технології, програмна реалізація інформаційної технології охорона праці та безпека в надзвичайних ситуаціях, висновки

5. Перелік графічного матеріалу (з точним зазначенням обов'язкових креслень)
Електронні плакати

6. Консультанти розділів проекту (роботи)

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
Охорона праці та безпека в надзвичайних ситуаціях	Критська Я.О. ст. викл. кафедри КНІ		

7. Дата видачі завдання 18.10.2018

Керівник

_____ (підпис)

Завдання прийняв до виконання

_____ (підпис)

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів дипломного проекту (роботи)	Строк виконання етапів проекту (роботи)	Примітка
1	Розробка технічного завдання	10.09.2018-15.09.2018	
2	Аналіз завдання та робота з літературою	16.09.2018-22.09.2018	
3	Розробка та програмна реалізація методу	26.09.2018-06.10.2018	
4	Тестування програмного забезпечення	07.10.2018-25.11.2018	
5	Розробка частини проекту "Охорона праці та безпеки в надзвичайних ситуаціях"	26.11.2018-1.12.2018	
6	Оформлення пояснювальної записки, автореферату та презентації	2.12.2018-09.01.2019	
7			

Студент

_____ (підпис)

Березін О.Є.

_____ (прізвище та ініціали)

Науковий керівник

_____ (підпис)

Недзельський Д.О.

_____ (прізвище та ініціали)

АНОТАЦІЯ

Березін О.Є. Методи та інформаційна технологія виявлення та класифікації двомірних об'єктів.

Магістерська робота присвячена дослідженню області розпізнавання двомірних об'єктів. Проведено аналіз предметної області, а також допоміжних знань, які можуть бути застосовні в цьому руслі. Розроблено та реалізовано алгоритм розпізнавання заданих плоских об'єктів, що працює в режимі реального часу і є незалежним від масштабів, зсувів, умов освітлення і поворотів в фокальній площині. Вдалося домогтися і певної поміхо стійкості, незалежності від часткових перекриттів самого об'єкта розпізнавання.

Ключові слова: метод, зображення, відеопоток, об'єкт, модель, дескриптор.

ABSTRACT

Berezin O.E. Methods and information technology for the identification and classification of two-dimensional objects.

The work deals with the field of recognition of flat objects. We conducted an analysis of the subject area as well as support of knowledge that can be applied in this way. Developed and implemented an algorithm for recognizing given planar objects, it works in real time and regardless of the size, displacement, lighting conditions and turns in the focal plane. It achieves a certain noise immunity, as well as independence from the partial overlap of the object recognition.

Keywords: method, image, video stream, object, model, descriptor.

АННОТАЦИЯ

Березин А.Е. Методы и информационная технология обнаружения и классификации двухмерных объектов.

Работа посвящена исследованию области распознавания плоских объектов. Проведён анализ предметной области, а также вспомогательных знаний, которые могут быть применимы в этом русле. Разработан и реализован алгоритм распознавания заданных плоских объектов, работающий в режиме реального времени и независимый от масштабов, смещений, условий освещения и поворотов в фокальной плоскости. Удалось добиться и определённой шумоустойчивости, а также независимости от частичных перекрытий самого объекта распознавания.

Ключевые слова: метод, изображения, видеопотоки, объект, модель, дескриптор.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ	5
ВСТУП.....	6
1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ	7
1.1 Постановка задачі дослідження	7
1.2 Огляд існуючих підходів	8
1.2.1 Контурне розпізнавання	8
1.2.2 Розпізнавання на основі ключових точок (surf)	16
1.3 Огляд варіантів застосування.....	31
2 РОЗРОБКА АЛГОРИТМУ ІНФОРМАЦІЙНОЇ ТЕХНОЛОГІЇ.....	39
2.1 Загальний план.....	39
2.2 Побудова моделі об'єкту	41
2.3 Обробка зображення «сцени»	43
2.4 Знаходження положення об'єкта на «сцені»	45
3 ПРОГРАМНА РЕАЛІЗАЦІЯ ІНФОРМАЦІЙНОЇ ТЕХНОЛОГІЇ.....	48
3.1 Розробка демо-додатку	48
3.2 Посібник користувача демо додатку.....	49
3.3 Аналіз результатів	53
4 ОХОРОНА ПРАЦІ ТА БЕЗПЕКА В НАДЗВИЧАЙНИХ СИТУАЦІЯХ.....	57
4.1 Аналіз потенційно небезпечних і шкідливих виробничих факторів , що впливають на персонал	57
4.2 Заходи щодо техніки безпеки.....	60
4.3 Заходи, що забезпечують виробничу санітарію і гігієну праці.....	63
4.4 Рекомендації по пожежній профілактиці.....	66
ВИСНОВКИ	69
ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАНЬ	71
ДОДАТОК А. Електронні плакати	73

**ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ,
СКОРОЧЕНЬ І ТЕРМІНІВ**

ЕОМ – електронно обчислювальна машина

ПК – персональний комп'ютер

НДЛ – науково-дослідна лабораторія

ЛМС – «Людина – машина - середина»

ВСТУП

У світі сучасних технологій наголос все більше і більше переходить від збільшення потужності апаратної частини до розробки розумного програмного забезпечення. У різних областях людської діяльності постійно виникають нові потреби в усіляких алгоритмах обробки даних і часто обсяг цих даних виявляється досить великий. У той же час на швидкодію накладається ряд вимог, який відсікає можливість використання інтуїтивних рішень. А в руслі комп'ютерного зору про інтуїтивні рішення взагалі можна забути. Тому нові алгоритми в області розпізнавання образів являють величезний інтерес.

У даній роботі мова піде про розпізнавання образів, а точніше про розпізнавання плоских об'єктів. Цей підрозділ примітний відсутністю об'ємних моментів і перетворень, яким об'єкти піддаються за рахунок своїх рухів щодо довільних осей, таких, як рух окремими частинами тіла у живих істот. Грубо кажучи, їх моделі можна описати в двовимірному просторі. Прикладами даних об'єктів можуть служити дорожні знаки, картини в музеї, зображення на сторінках книг, різні коди, нанесені на ту чи іншу продукцію, тощо.

Розпізнавання плоских об'єктів - це окремий випадок розпізнавання образів і тут можливо було б застосувати алгоритми машинного зору, мета у яких, набагато більше, ніж плоскі об'єкти. До таких можна віднести побудову каскадів класифікаторів і дерев рішень. Однак відмітні властивості плоских об'єктів говорять нам про те, що можливо існують більш ефективні методи вирішення цієї проблеми.

В рамках цієї роботи будуть розглянуті різні підходи до розпізнавання плоских об'єктів, проте найбільша увага приділятиметься роботі з ключовими точками і характерними ознаками зображень, які отримані алгоритмом SURF. Будуть розглянуті методи машинного навчання стосовно цих характеристик, методи їх співставлення та фільтрації, а також подальшої обробки.

Також торкнемося принципових обмежень, які накладаються на плоскі об'єкти і допомагають при їх подальшому розпізнаванні.

У процесі виконання дипломної роботи також було проаналізовано умови праці згідно з потребами розділу охорони праці.

1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ

1.1 Постановка задачі дослідження

Людина може порівнювати зображення і виділяти на них об'єкти візуально, на інтуїтивному рівні. Однак, для машини зображення - всього лише набір даних, який ні про що не говорить. У даній роботі піде мова про те як «наділити» машину умінням розпізнавати наперед задані плоскі об'єкти по їх візуальному опису.

Подання об'єкта для розпізнавання відбувається подачею зображення, на якому представлений цей об'єкт. Це зображення назвемо еталоном. Ґрунтуючись на ньому, будуть проводитися подальші пошуки. Шукати це зображення необхідно на іншому зображенні або на відео-потоці. Відео-потік в свою чергу є набором зображень. Тобто, завдання зводиться до знаходження одного зображення на іншому.

Перелічимо деякі перетворення, щодо яких ми б хотіли отримати інваріантність:

- 1) зміщення;
- 2) поворот в фокальній площині;
- 3) зміна масштабу;
- 4) зміни яскравості;
- 5) зміни положення камери / поворот об'єкта у довільних напрямках.

Інваріантність щодо трьох останніх перетворень в повній мірі отримати не вдасться, але хоча б частково зробити це вже було б не погано.

Також треба отримати допустиму поміхостійкість і незалежність від фону, щоб алгоритм був застосовний у реальних ситуаціях.

У плані швидкодії потребується можливість обробки відео-потoku у реальному часі, тобто $24 \pm$ кадрів (640×480) на секунду.

Метою магістерської роботи є дослідження і вдосконалення попередніх методів виявлення та класифікації двомірних об'єктів.

В якості об'єкта дослідження розглянуто інформаційну технологію, що виконує виявлення двомірного об'єкта як в статичному зображенні, так і у відео потоці.

Для досягнення поставленої мети необхідно виконати:

- дослідження і аналіз існуючих методів розпізнавання об'єкта;
- розробку методу виявлення та розпізнавання об'єктів у відеопотоці;
- розробку методу аналізу та фільтрації ключових точок зображення об'єкту.

1.2 Огляд існуючих підходів

Є певні методи для порівняння зображень, засновані на зіставленні знань про зображення в цілому. У загальному випадку це виглядає наступним чином: для кожної точки зображення обчислюється значення певної функції, на підставі цих значень можна приписати зображенню певну характеристику, тоді порівняння зображень зводиться до задачі порівняння таких характеристик. Ці методи настільки ж погані, наскільки прості, і працюють прийнятно, практично, тільки в ідеальних ситуаціях. Причин цьому предостатньо: поява нових об'єктів на зображенні, перекриття одних об'єктів іншими, шуми, зміни масштабу, положення об'єкта на зображенні, положення камери в тривимірному просторі, освітлення, афінні перетворення і т. п. Власне, погані якості цих методів обумовлені їх основною ідеєю, тобто тим, що в характеристику вносить вклад кожна точка зображення, яким би поганим цей внесок не був.

Остання фраза відразу наштовхує на думки обходу таких проблем: треба або якось вибрати точки, які роблять внесок у характеристику, або, ще краще, виділяти деякі особливі (ключові) ознаки і порівнювати їх. Зараз виділяють два основних види таких ознак: контури і ключові точки. Вони досить сильно різняться і в основному використовуються для розпізнавання різних видів зображень.

1.2.1 Контурне розпізнавання

Якщо ж ми самі створюємо зображення, яке плануємо знайти надалі, то у нас є можливість намалювати його таким чином, що б надалі легко розпізнавати. Тут у справу можуть вступити контурні ознаки.

На практиці ми обмежені дозволом веб-камери (телефону), особливостями передачі кольору, освітлення і обчислювальною потужністю обладнання (все ж відбувається в реальному часі, а тому має робитися швидко), а тому вибирається зазвичай чорно-білий маркер (еталонне зображення) простої форми. Як правило, це прямокутник або квадрат з вписаним у всередину ідентифікатором. Можливі види маркерів наведені на рисунках 1.1 – 1.4.

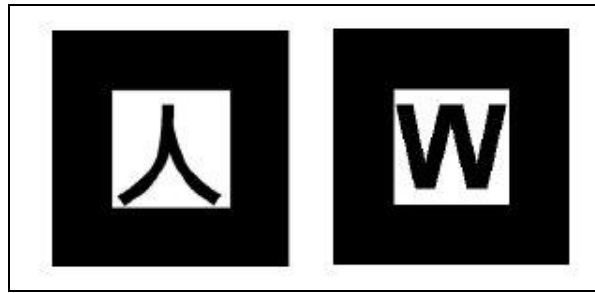


Рисунок 1.1 – ArToolKit(ATK) marker system

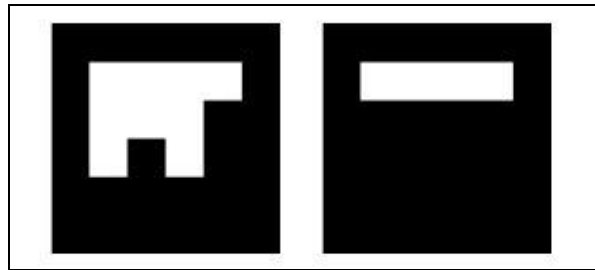


Рисунок 1.2 – Institut Graphische Datenverarbeitung (IGD) marker system

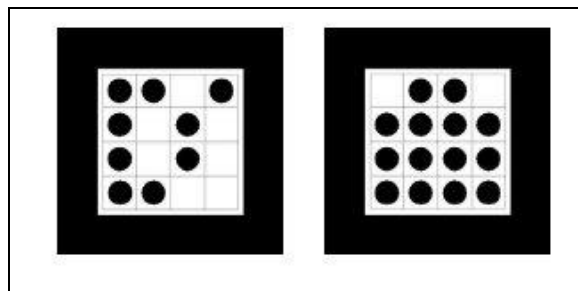


Рисунок 1.3 – Siemens Corporate Research (SCR) marker system

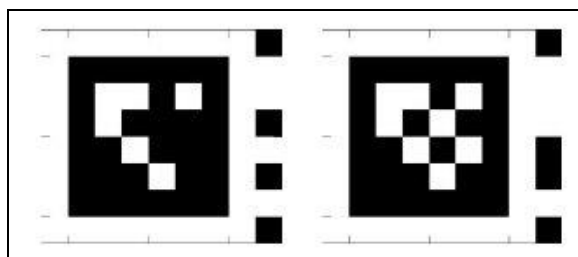


Рисунок 1.4 – Hoffman marker system (HOM)

Якщо коротко, алгоритм розпізнавання маркера виглядає так:

- 1) Переводимо в градації сірого кольору;
- 2) Бінаризація зображення (поріг);
- 3) Визначення замкнутих областей;
- 4) Виділяємо контури;
- 5) Виділяємо кути маркера;

б) Перетворюємо координати.

Поетапна обробка зображення показана на рисунку 1.5.

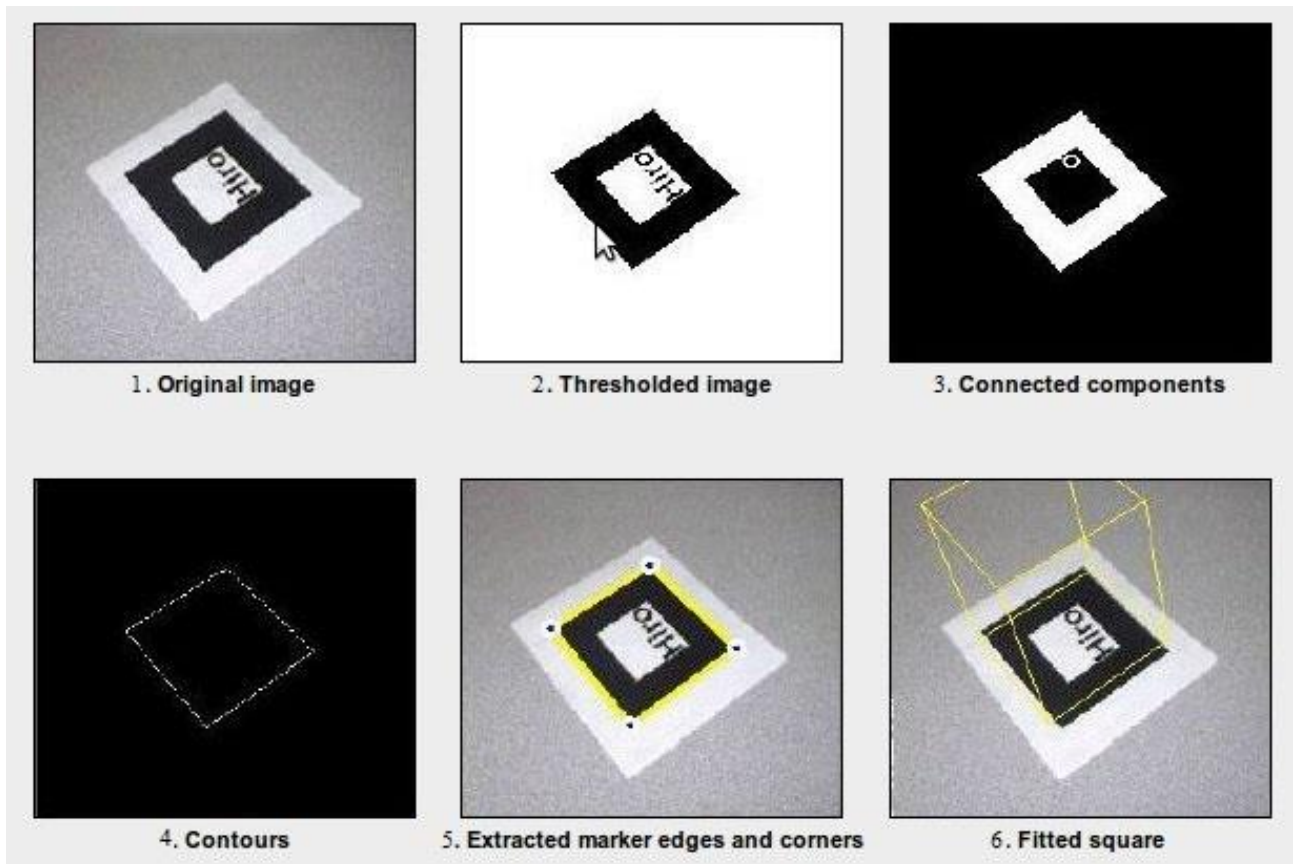


Рисунок 1.5 – Етапи контурного розпізнавання маркера

Розглянемо більш докладно, що відбувається на кожному кроці з ілюстраціями.

Наведемо три алгоритму переведення кольорового зображення в градації сірого.

1) Світлота (Lightness): $GS = (\max(R,G,B) + \min(R,G,B))/2$,

2) Світність (Luminosity): $GS = 0.21 \times R + 0.72 \times G + 0.07 \times B$,

3) Середнє (Average): $GS = (R + G + B) / 3$.

Приклади, як виглядають ці три способи наведені на рисунку 1.6. Як правило використовують "Світність".

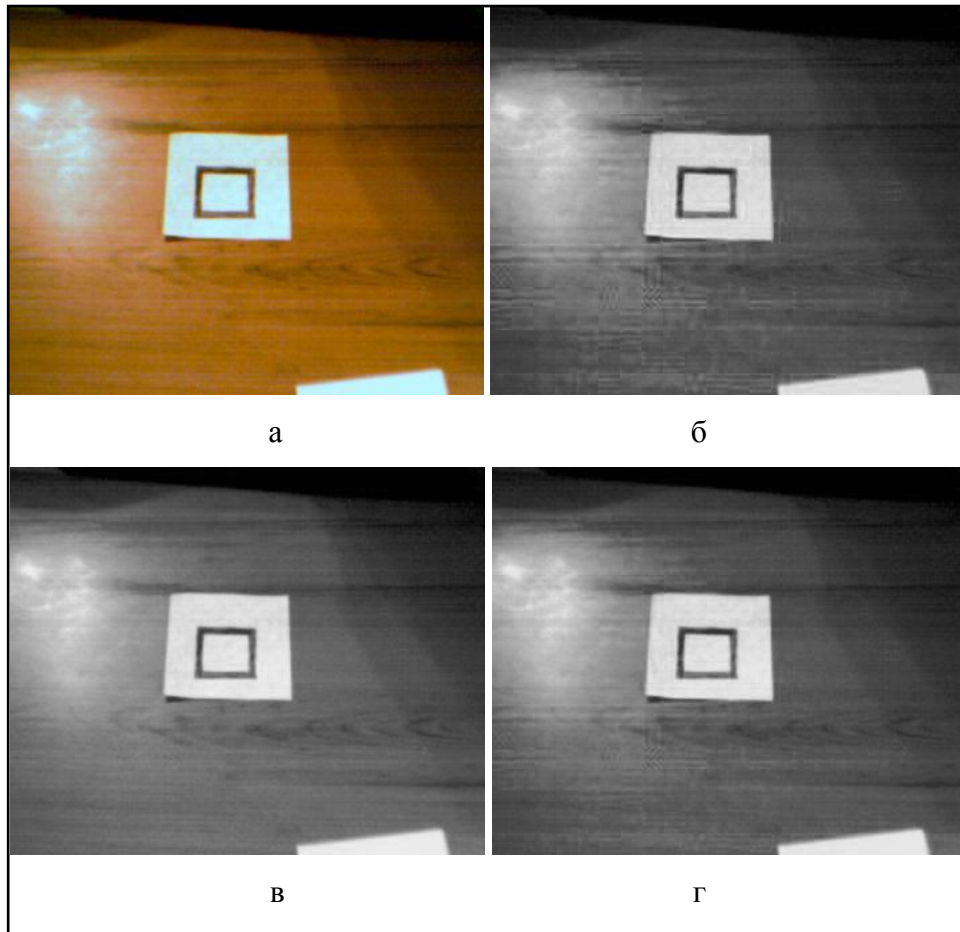


Рисунок 1.6 – результати роботи різних алгоритмів перетворення в ч / б:
а - вихідне зображення, б - світлота, в - світність, г - середнє

Для переведення зображення в двокольоровий стан використовується певний поріг. Питання в тому, як і до чого цей поріг застосовувати. Найпростіший спосіб - задати поріг. Наприклад, якщо у нас 256 кольорів, то можемо задати поріг 128. Якщо трохи ускладнити - то можемо вибрати поріг, користуючись гістограмою кольору. Взагалі всі методи перетворення зображення до бінарного вигляду можна розділити на шість великих груп які перераховані і класифіковані в статті.

- 1) Методи, засновані на "формі" гістограми;
- 2) Методи на основі кластеризації;
- 3) Методи на основі вивчення ентропії;
- 4) Методи, засновані на пошуку подібності між сірим і ч / б зображенням;
- 5) Методи, що використовують кореляційні залежності та особливості статистичного розподілу між пікселями в областях зображення;
- 6) Методи, засновані на локальній адаптації порога для кожного пікселя зображення.

Найчастіше використовуються методи на основі локальної адаптації. Методів бінаризації близько сорока, так що наведемо приклад бінаризації зображення із заданим порогом, адаптивного, а також методу Otsu.



Рисунок 1.7 – Бінаризація зображення (жорстко заданий поріг)

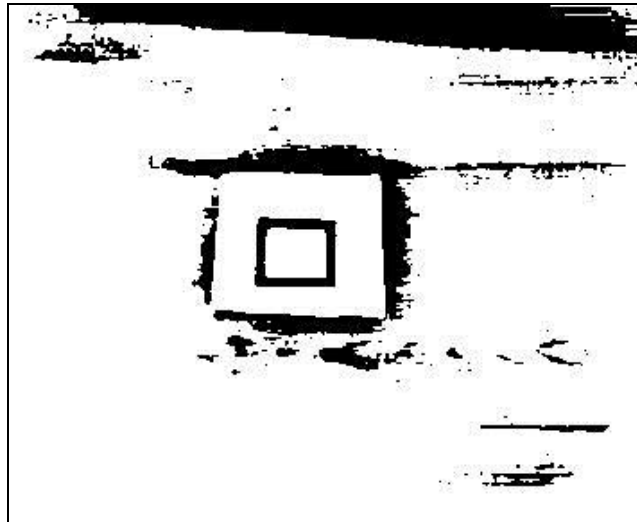


Рисунок 1.8 – Бінаризація зображення (адаптивна бінаризація)

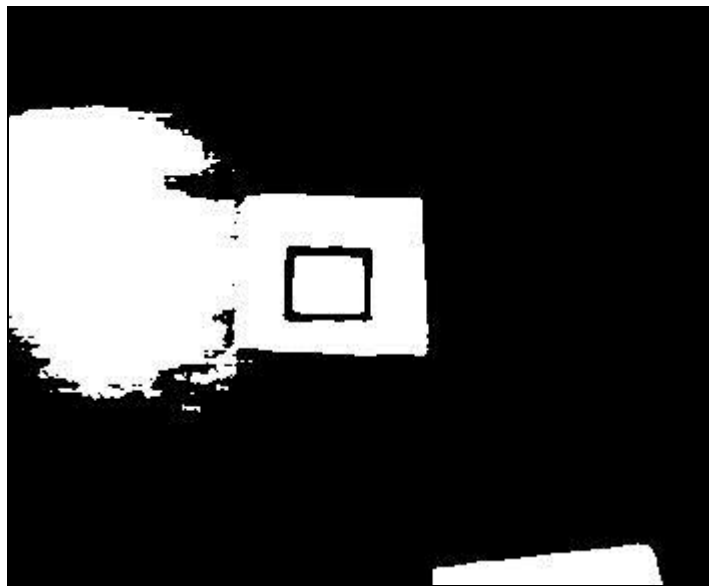


Рисунок 1.9 – Бінаризація зображення (метод Otsu)

Нам необхідно визначити замкнуті області на білому фоні. Тут йде комбінація алгоритмів, але в загальному випадку - як правило застосовують алгоритми "заливки" білих областей і виділяють замкнуті області.

У цьому місці можливі варіації. Спочатку зображення перевіряється на замкнутість, а потім виділяються контури (рис. 1.10) або спочатку виділяють контури, а потім перевіряють на замкнутість (рис. 1.11).

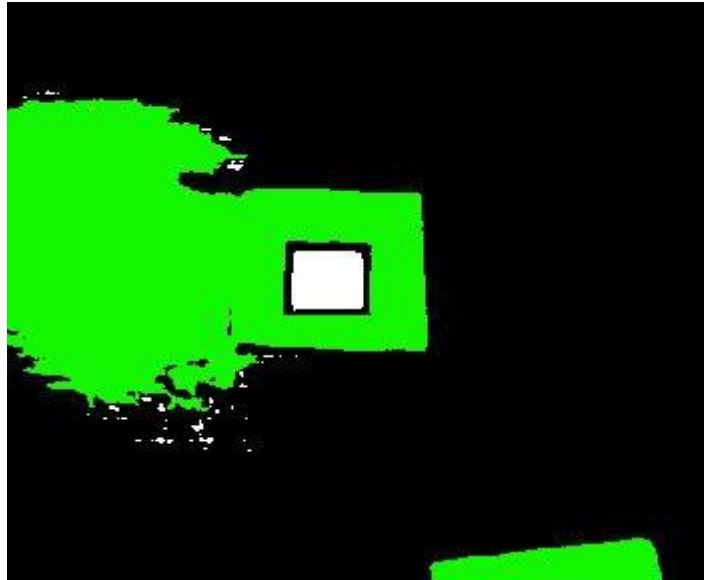


Рисунок 1.10 – Перевірка на замкнутість, потім виділення контуру

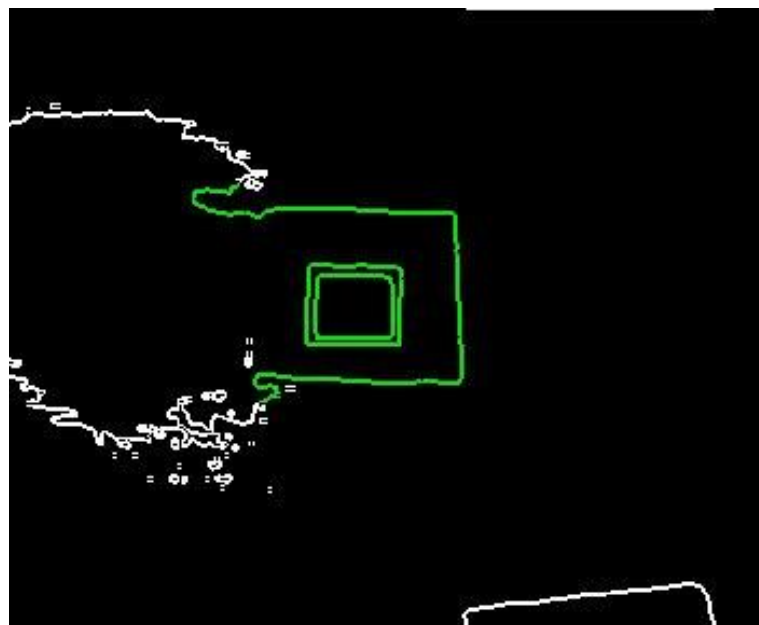


Рисунок 1.11 – Виділення контуру, потім перевірка на замкнутість

Існує кілька підходів до виділення контурів на зображенні. Перерахуємо лише основні:

- 1) Marr-Hildreth Edge Detector;

- 2) Canny Edge Detector;
- 3) Boolean function based Edge Detector;
- 4) Euclidian distance and Vector Angle based Edge Detector;
- 5) Depth Edge Detection using Multi-Flash Imaging;
- 6) Sobel Edge Detector.

Знову-таки, матеріал досить об'ємний, порівняння та опис алгоритмів може зайняти не одну сторінку. В основному використовується алгоритм Канні (рис. 1.12) і Собеля (рис. 1.13). Всередині білих замкнутих областей шукаємо контури.

В результаті застосування, отримуємо контур ґрунтуючись на якому будемо виділяти положення маркера.

У детекторі кордонів Канні є варіативність з використанням шляхів завдання різних параметрів.

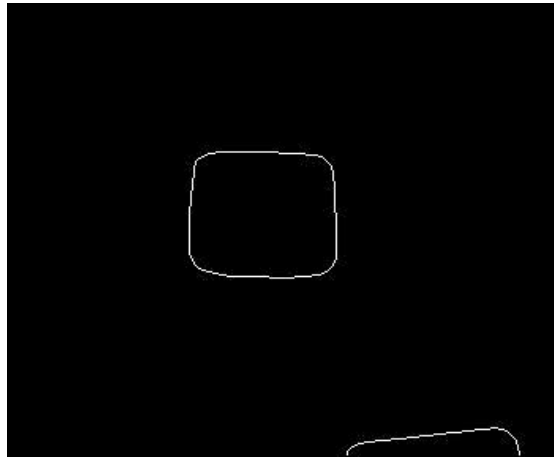


Рисунок 1.12 – Результат виділення контурів детектором Канні

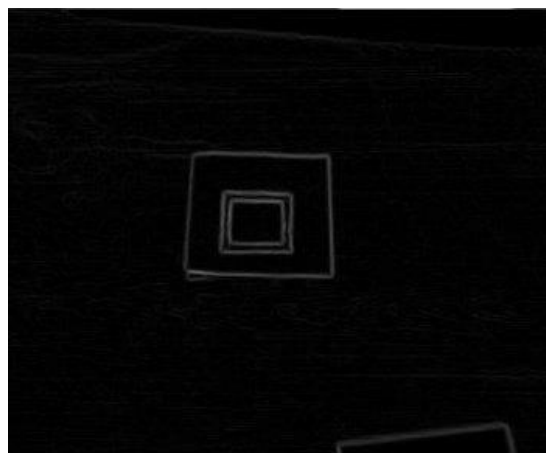


Рисунок 1.13 – Результат виділення контурів детектором Собеля

Виділивши контур, його необхідно зіставити з нашим маркером. Може бути виділено багато різних контурів на зображенні, нам необхідно знайти щось «схоже» на чотирикутник.

Для вирішення цього завдання можна застосувати алгоритм Дугласа-Пекера (він же алгоритм Рамер-Дугласа-Пекера, алгоритм ітеративної найближчої точки, алгоритм розбиття і злиття), що дозволяє зменшити кількість точок кривої, апроксимувати великою серією точок.

Наприклад, в бібліотеці OpenCV є функція `approxPolyDP`, яка вже це робить. Якщо її результат додатково обробити, то вийде цілком підходящий результат.

Таким чином, ми отримали координати кутів маркера.

Тепер у нас є координати кутів маркера, які, по суті, в ідеалі є перпендикулярними, а в реальності розташовані під іншим кутом. Крім того, і в ідеалі і в реальності, сторони квадрата є осями координат. Таким чином, ми можемо визначити положення «камери» щодо нашого об'єкта, і точку відліку початку координат.

На чому заснований метод визначення координат можна показати графічно (рис. 1.14).

Ідея полягає в тому, що при зміні кута, з якого дивиться камера, змінюється розмір проекції. Наприклад, знаючи положення камери і точки відліку, ми можемо намалювати проекцію для нашої 3D моделі. В даному випадку куба.

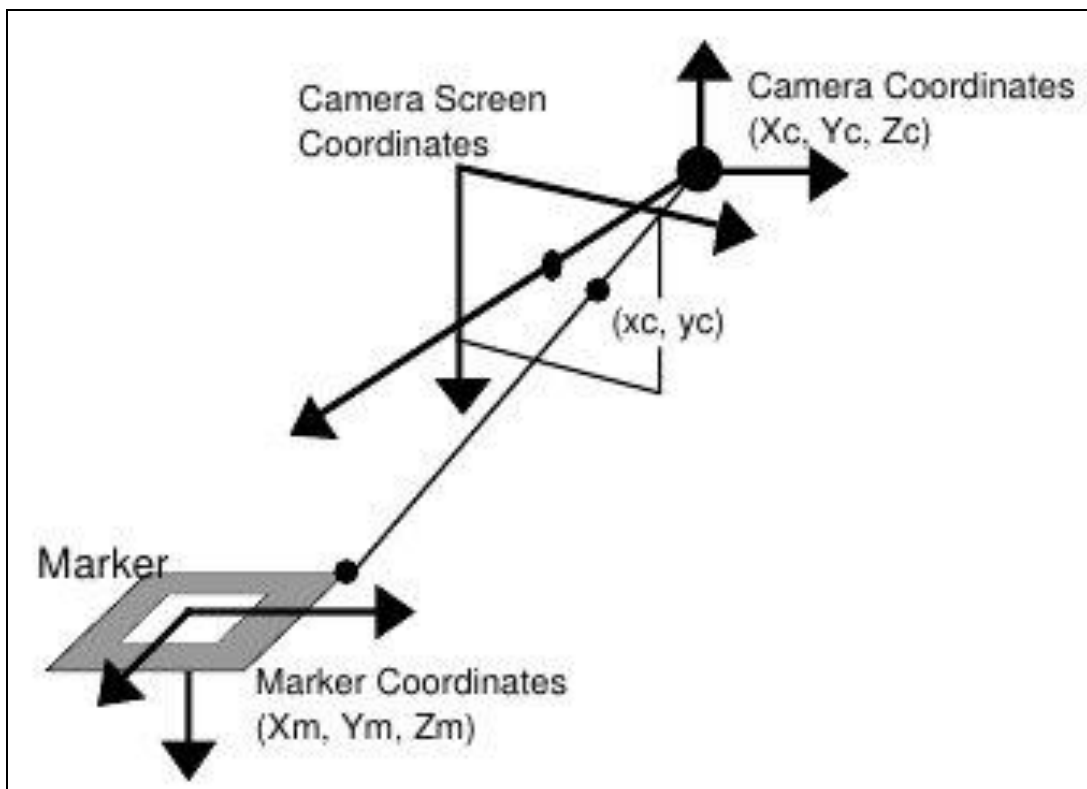


Рисунок 1.14 – Визначення координат маркера

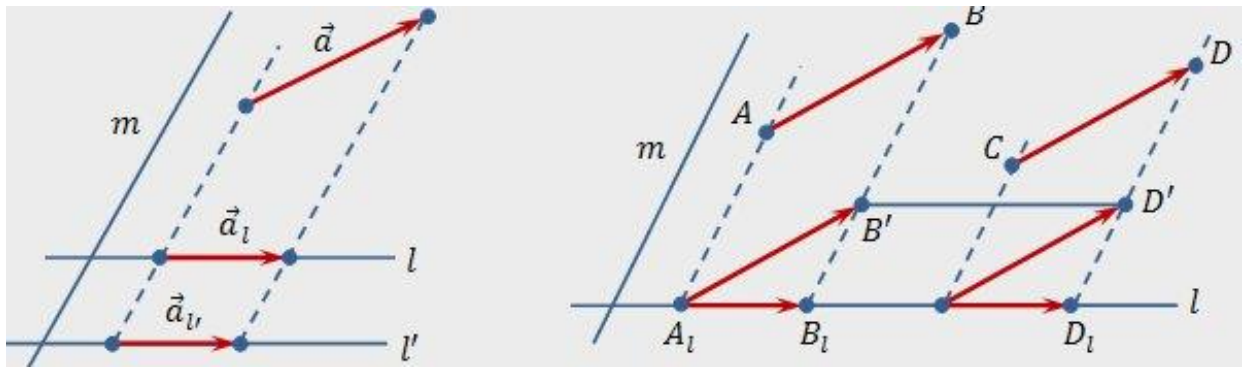


Рисунок 1.15 – Проекція 3D моделі

Як бачимо, якщо ми використовуємо квадрат з чистим полем в якості маркера, то він симетричний, і розпізнати обертання ми можемо тільки лише частково. Іноді цього достатньо. Якщо ж ні - то ми можемо внести всередину квадрата додатковий маркер (рис. 1.16). Потім отримати в результаті ще й кут повороту, використовуючи перетворення Хаффа.



Рисунок 1.16 – Вкладений маркер

1.2.2 Розпізнавання на основі ключових точок (surf)

Ми і підійшли до ідеї зіставлення зображень по ключових точках. Можна сказати, що ми замінюємо зображення деякої моделлю - набором його ключових точок. Відразу відзначимо, що особливою буде називатися така точка зображеного об'єкта, яка з великою часткою ймовірності буде знайдена на іншому зображенні цього ж об'єкта. Детектором будемо називати метод вилучення ключових точок з зображення. Детектор повинен забезпечувати інваріантність знаходження одних і тих же особливих точок щодо перетворень зображень.

Єдине, що залишається незрозумілим - яким чином визначати яка ключова точка одного зображення відповідає ключовій точці іншого зображення. Адже після застосування

детектора можна визначити тільки координати особливих точок, а вони на кожному зображенні різні. Тут у справу і вступають дескриптори. Дескриптор - ідентифікатор ключової точки, що виділяє її з іншої маси особливих точок. У свою чергу, дескриптори повинні забезпечувати інваріантність знаходження відповідності між особливими точками щодо перетворень зображень.

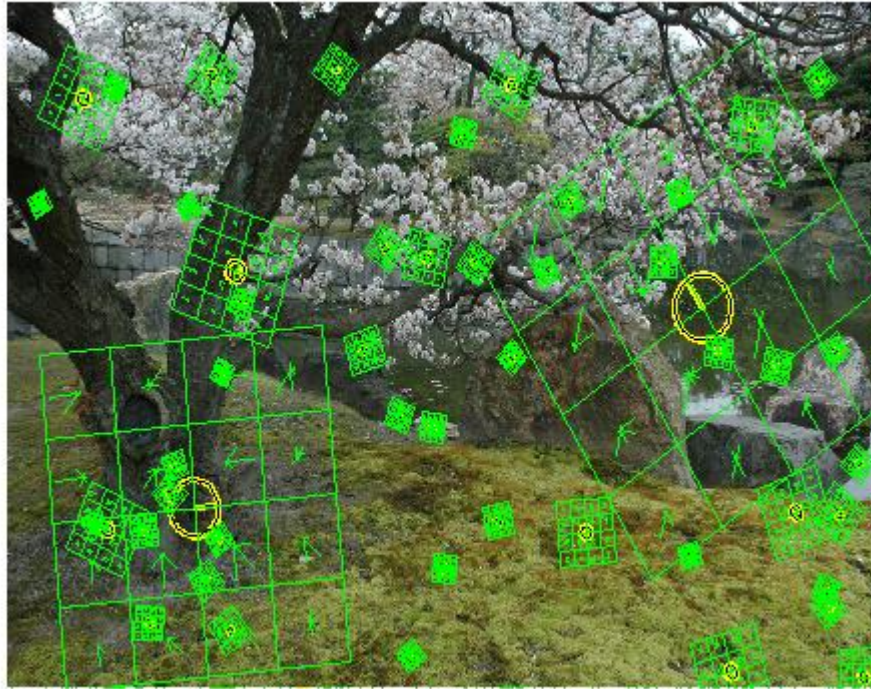


Рисунок 1.17 – Приклад ключових точок зображення і їх дескрипторів

У підсумку виходить наступна схема рішення задачі зіставлення зображень:

- 1) На зображеннях виділяються ключові точки і їх дескриптори;
- 2) За збігом дескрипторів виділяються відповідні один одному ключові точки;
- 3) На основі набору ключових точок, що збіглися, будується модель перетворення зображень, за допомогою якого з одного зображення можна отримати інше.

На кожному етапі є свої проблеми і різні методи їх вирішення, що вносить певне свавілля в рішення початкового завдання. Так існує не один метод для розпізнавання на основі ключових точок.

Серед різних споріднених методів, розглянемо метод Speeded Up Robust Features (SURF), оскільки він є одним з найбільш ефективних і швидких сучасних алгоритмів. Крім того, SURF є поширеним методом, його реалізації є в багатьох математичних бібліотеках.

SURF вирішує два завдання - пошук особливих точок зображення і створення їх дескрипторів, інваріантних до масштабу і обертанню. Це означає, що опис ключової точки буде однаковий, навіть якщо зразок змінить розмір і буде повернений (тут і далі ми будемо говорити тільки про обертання в площині зображення - фокальній площині). Крім того, сам

пошук ключових точок теж повинен володіти інваріантністю. Так, що б повернений об'єкт сцени мав той же набір ключових точок, що і зразок.

Метод шукає особливі точки за допомогою матриці Гессе. Детермінант матриці Гессе (т.зв. гессіан) досягає екстремуму в точках максимальної зміни градієнта яскравості. Він добре детектує плями, кути і краї ліній.

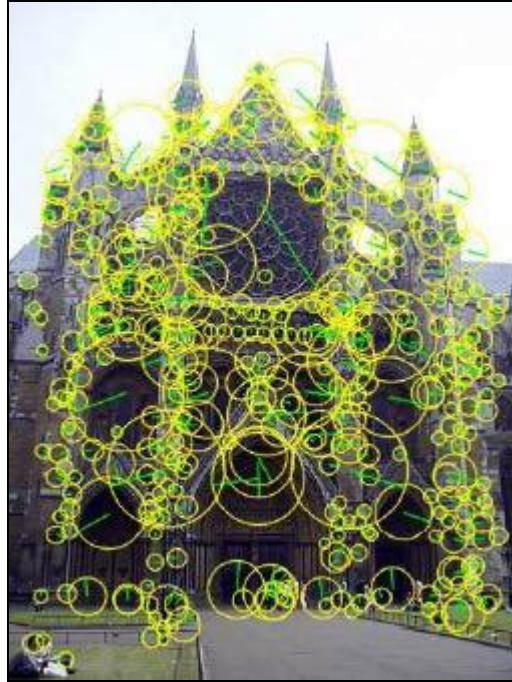


Рисунок 1.18 – Особливі точки зображення будівлі, знайдені за допомогою матриці Гессе. Діаметр кола показує масштаб особливої точки. Зелена лінія - напрям градієнта яскравості

Гессіан інваріантний щодо обертання. Але не інваріантний до масштабу. Тому SURF використовує різномасштабні фільтри для знаходження гессіанов.

Для кожної ключової точки розраховується напрям максимальної зміни яскравості (градієнт) і масштабу, взятий з масштабного коефіцієнта матриці Гессе. Градієнт в точці обчислюється за допомогою фільтрів Хаара.

Після знаходження ключових точок, SURF формує їх дескриптори. Дескриптор являє собою набір з 64 (або 128) чисел для кожної ключової точки. Ці числа відображають флуктуації градієнта навколо ключової точки (що розуміється під флуктуацією - розглянемо нижче). Оскільки ключова точка являє собою максимум гессіана, то це гарантує, що в околиці точки повинні бути ділянки з різними градієнтами. Таким чином, забезпечується дисперсія (відмінність) дескрипторів для різних ключових точок.

Флуктуації градієнта околиць ключової точки розраховуються відносно напрямку градієнта навколо точки в цілому (по всій околиці ключової точки). Таким чином, досягається інваріантність дескриптора щодо обертання. Розмір же області, на якій

розраховується дескриптор, визначається масштабом матриці Гессе, що забезпечує інваріантність щодо масштабу.

Флуктуації градієнта також розраховуються за допомогою фільтра Хаара.

Для ефективного обчислення фільтрів Гессе і Хаара - використовується інтегральне представлення зображень.

Якщо коротко, то інтегральне представлення є матрицею, розмірність якої збігається з розмірністю вихідного зображення, а елементи розраховуються за формулою:

$$I(x, y) = \sum_{i=0, j=0}^{i \leq x, j \leq y} I(i, j), \quad (1.1)$$

де $I(i, j)$ - яскравість пікселів вихідного зображення.

Маючи інтегральну матрицю можна дуже швидко обчислювати суму яскравості пікселів довільних прямокутних областей зображення, за формулою:

$$S(ABCD) = I(A) + I(C) - I(B) - I(D), \quad (1.2)$$

де ABCD – заданий прямокутник.

Виявлення особливих точок в SURF засновано на обчисленні детермінанта матриці Гессе (гессіан).

Матриця Гессе для двовимірної функції та її детермінант визначається таким чином:

$$H(f(x, y)) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} \quad (1.3)$$

$$\det(H) = \frac{\partial^2 f}{\partial x^2} \frac{\partial^2 f}{\partial y^2} - \left(\frac{\partial^2 f}{\partial x \partial y} \right)^2,$$

для знаходження локального мінімуму або максимуму яскравості зображення. У цих точках значення гессіана досягає екстремуму.

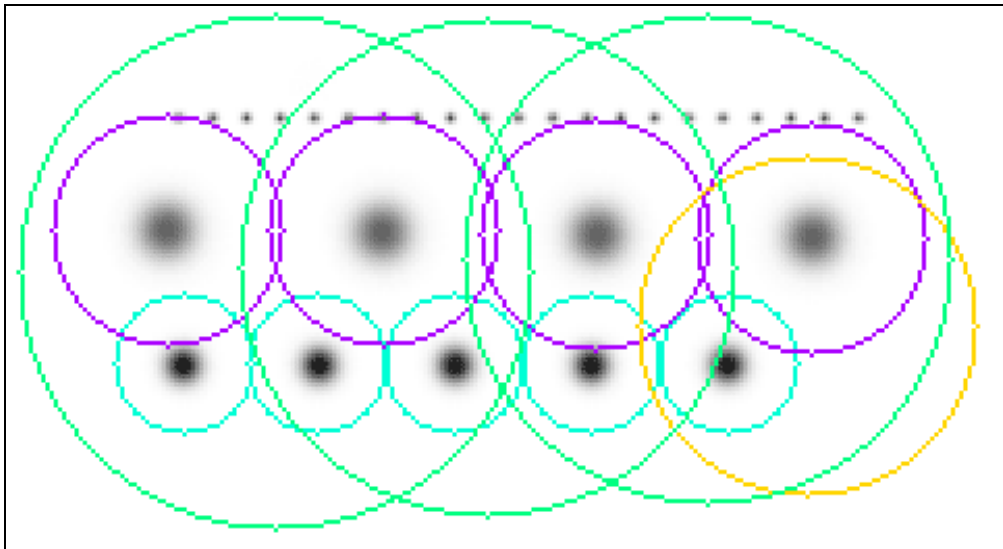


Рисунок 1.19 – Особливі точки (окреслені кольоровими колами) являють собою локальні екстремуми яскравості зображення. Дрібні крапки не розпізнані як особливі, через порогові відсікання за величиною гессіана

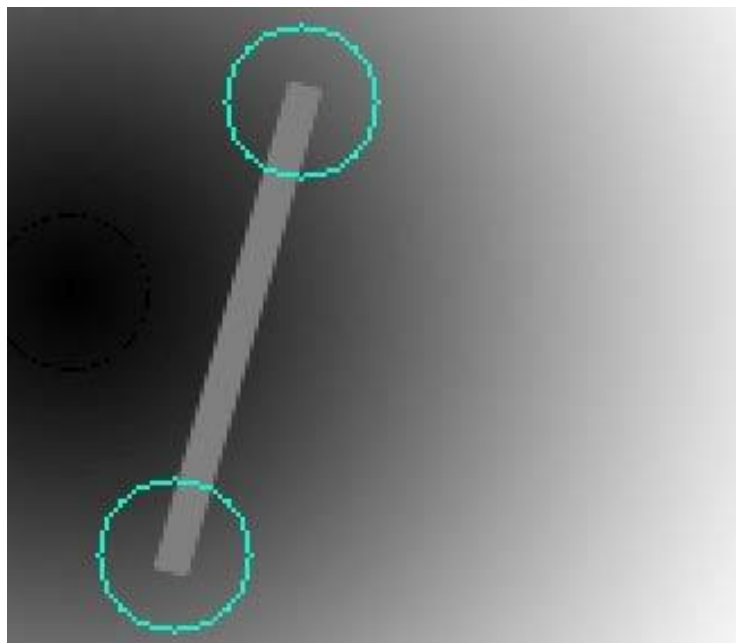


Рисунок 1.20 – Кінці відрізка, розпізнані як ключові точки, за допомогою матриці Гессе

Теоретично, обчислення матриці Гессе зводиться до знаходження Лапласіан Гауссіан. По суті, елементи матриці Гессе обчислюються як згортка (сума добутків) пікселів зображення на фільтри, зображені на рисунку 1.21.

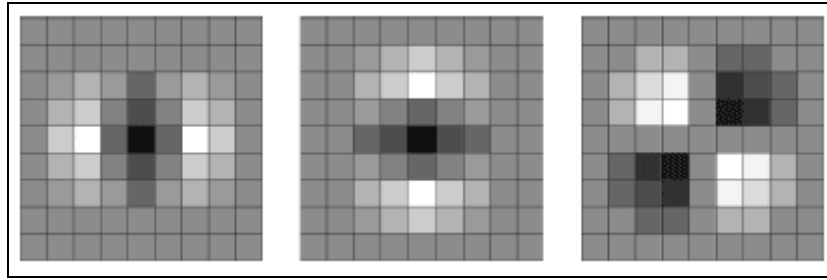


Рисунок 1.21 – Фільтри матриці Гессе

На рисунку 1.21 зображені дискретизовані фільтри для знаходження чотирьох елементів матриці Гессе (четвертий - збігається з третім, оскільки матриця Гессе симетрична). Фільтри мають просторовий масштаб 9x9 пікселів. Темні ділянки відповідають негативним значенням фільтра, світлі - позитивним.

Однак, SURF не використовує лапласіан гауссіана в тому вигляді, який зображений на малюнку. По-перше, дискретизований лапласіан гауссіана має досить великий розкид значення детермінанта, при обертанні зразка (нагадаємо, що в ідеалі гессіан повинен бути інваріантний до обертання). Особливо детермінант «просідає» в районі повороту на 45 градусів. А по-друге, і це головне, фільтр для Лапласіан гауссіана має безперервний характер. Майже всі пікселі фільтра мають різні величини яскравості. А це не дозволяє ефективно використовувати такий потужний механізм розрахунку, як інтегральну матрицю зображення.

Тому SURF використовує бінарізовану апроксимацію Лапласіан Гауссіан (автори назвали його Fast-Hessian – рис. 1.22).

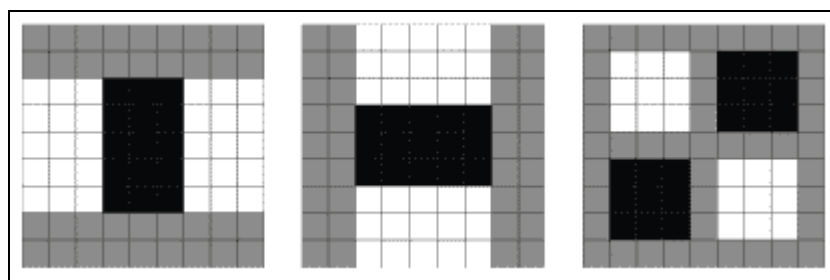


Рисунок 1.22 – Модифіковані фільтри матриці Гессе застосовувані в алгоритмі SURF. Fast-Hessian

На рисунку 1.22 зображені фільтри, використовувані для знаходження матриці Гессе в SURF. Білі області відповідають значенням +1, чорні -2 (на третьому фільтрі -1), сірі - нульові. Просторовий масштаб - 9x9 пікселів.

Цей фільтр більш стійкий до обертання, і його можна ефективно обчислити за допомогою інтегральної матриці.

Таким чином, в SURF, гессіан обчислюється так:

$$\det(H_{approx}) = D_{xx}D_{yy} - (0.9D_{xy})^2 \quad (1.4)$$

де D_{xx} , D_{yy} , D_{xy} - згортки по фільтрам, зображеним на рисунку 1.22. Коефіцієнт 0.9 має теоретичне обґрунтування, і коригує наближений характер обчислень.

Отже, для знаходження особливих точок, SURF пробігається по пікселям зображення і шукає максимум гессіана. Спосіб знаходження локального максимуму гессіана ми розглянемо пізніше. У методі задається граничне значення гессіана. Якщо обчислене значення для пікселя вище порога - піксель розглядається як кандидат на ключову точку.

Тут ще корисно зауважити наступне. Оскільки гессіан є похідною, і залежить тільки від перепаду яскравості, але не від абсолютного її рівня, то він інваріантний по відношенню до зрушення яскравості зображення. Таким чином, зміна рівня освітлення зразка не впливає на виявлення ключових точок.

Крім того, властивості гессіана такі, що він досягає максимуму, як в точці білої плями на чорному тлі, так і чорної плями на білому тлі. Таким чином, метод виявляє і темні, і світлі особливості зображення.

Приклад пошуку ключових точок алгоритмом SURF, що відображає інваріантність стосовно пошуку білого на чорному або чорного на білому зображений на рисунку 1.23.

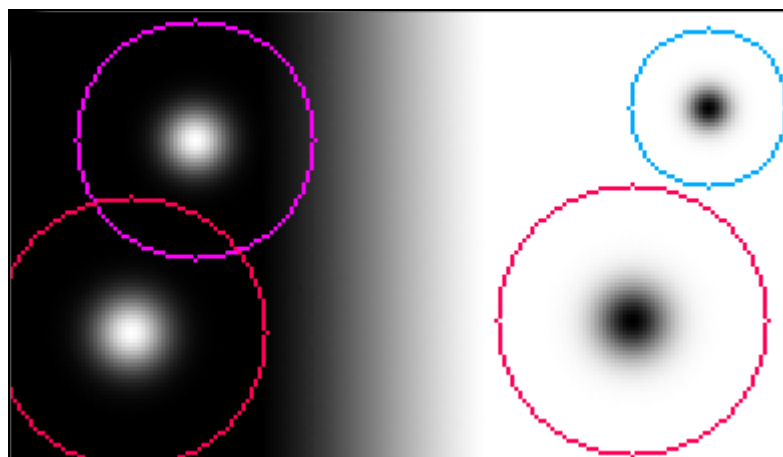


Рисунок 1.23 – світлі точки на темному фоні і темні точки на світлому фоні

Як уже зазначалося, гессіан не інваріантний щодо масштабу. Це означає, що для одного і того ж пікселя, гессіан може мінятися при зміні масштабу фільтра. Вирішення цієї

проблеми тільки одне - перебирати різні масштаби фільтрів і по черзі їх застосовувати до даного пікселю.

З міркувань симетрії і дискретизації, розмір фільтра Fast-Hessian не може приймати довільні значення. Допустимі розміри цього фільтра такі (починаючи з мінімального): 9, 15, 21, 27 і так далі, з кроком 6. Однак, на практиці, поступово збільшувати розмір фільтра на 6 - не вигідно, тому що для великих масштабів крок 6 виявляється занадто дрібним, а фільтри - надлишковими. Тому (і з деяких інших причин), SURF розбиває всі безліч масштабів на так звані октави. Кожна октава покриває певний інтервал масштабів, і має свій характерний розмір фільтра.

При цьому якби на октаву припадав тільки один фільтр, це було б занадто грубим наближенням. Крім того, ми б не могли знайти локальний максимум гессіана, серед різних масштабів, у різних октавах. Адже одна і та ж точка може мати кілька локальних максимумів гессіана, в різних масштабах. Це добре видно на рисунку 1.24.

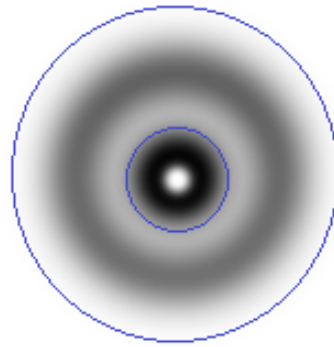


Рисунок 1.24 – Дві ключові точки різного масштабу в одній точці зображення

Якщо ми будемо шукати максимум серед усіх гессіанов, за всіма масштабами, то ми б знайшли тільки один з максимумів, в той час як їх може бути кілька. Один - в одному масштабі, інший - в іншому.

Виходячи з перерахованого, октава містить не один фільтр, а чотири фільтра, які добре покривають характерний масштаб октави (рис. 1.25).

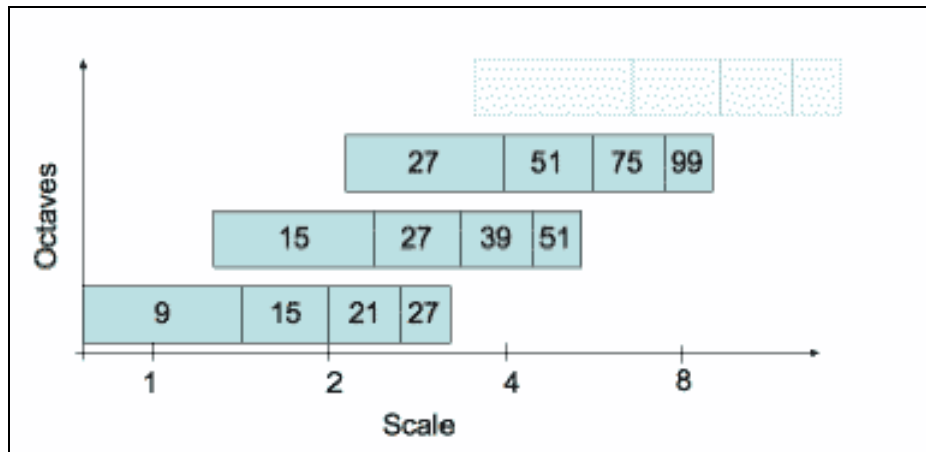


Рисунок 1.25 – групування різних масштабів в окремі октави

На рисунку 1.25 показані перші три октави методу SURF. Цифри в прямокутниках показують розмір фільтра Fast-Hessian. Логарифмічна шкала знизу - показує масштаби, що покриваються октавами.

Крок розміру фільтра в першій октаві - складає 6, у другій - 12, у третій - 24, у четвертій - 48 і так далі.

Як бачимо, октави значно перекриваються один одним. Це збільшує надійність знаходження локальних максимумів. Чому в октаві саме чотири фільтра стане ясно з наступного розділу.

Виникає питання, а скільки власне октав достатньо для покриття безлічі особливих точок різних масштабів? Теоретично, масштаби нескінченні, проте в реальних зображеннях, вони цілком кінцеві, і основна маса зосереджена в інтервалі від 1 до 10 (за даними авторів методу). Для покриття цього діапазону достатньо чотирьох октав. Плюс додається одна або дві октави для покриття більших масштабів. Разом, використовується 5-6 октав. Теоретично, цього цілком достатньо для покриття всіляких масштабів на зображенні 1024x768 пікселів.

Для знаходження локального максимуму гессіана, використовується так званий метод сусідніх точок 3x3x3.

Його сенс зрозумілий з рисунка 1.26.

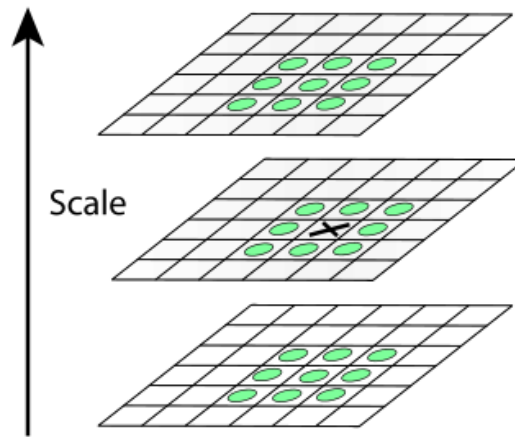


Рисунок 1.26 – Визначення точки локального максимуму або мінімуму

На рисунку 1.26 піксель, позначений хрестиком вважається локальним максимумом, якщо його гессіан більше ніж у будь-якого його сусіда в його масштабі, а також більше будь-якого із сусідів масштабом менше і масштабом більше (всього 26 сусідів).

Виходячи з такого визначення локального максимуму, зрозуміло, що октава повинна містити не менше трьох фільтрів, інакше ми не зможемо визначити факт знаходження локального максимуму гессіана всередині октави.

Відзначимо ще такий момент. Фільтри октави розраховуються не для всіх пікселів поспіль. Перша октава розраховується для кожного другого пікселя зображення. Друга - для кожного четвертого, третя - для кожного восьмого і так далі. Сенс зрозумілий - дві точки з відстанню 2 не можуть містити більше одного максимуму масштабу 2, 3 або більш високих масштабів. Тому немає сенсу перебирати всі точки зображення, для знаходження максимуму масштабу 3, наприклад.

Подвоєння кроку пікселів для октав дозволяє економити при розрахунку фільтрів. Як ви напевно вже помітили, розміри фільтрів в октавах повторюються. Так, наприклад, фільтр розміром 27 присутній в трьох октавах. Так от, при обчисленнях, цей фільтр буде розраховуватися тільки для першої октави. Друга і третя - просто використовують розрахунки першої октави. А подвоєння кроку пікселів гарантує, що точки в яких потрібно розрахувати гессіан, вже були прораховані попередньою октавою.

Тому, незважаючи на те, що октава містить чотири фільтра, насправді кожна октава (крім першої) розраховує тільки два характерних для неї розміру, два інших - завжди можна взяти з попередніх октав. Перша ж октава змушена розрахувати всі чотири своїх фільтра.

Отже, після знаходження максимального гессіана методом сусідніх точок 3×3 , ми знайшли піксель, в якому цей максимум досягається. Однак, оскільки, октава перебирає не всі точки зображення, то істинний максимум може не збігатися з знайденим пікселем, а лежати десь поруч, в сусідніх пікселях.

Для знаходження точки істинного максимуму, використовується інтерполювання знайдених гессіанов куба $3 \times 3 \times 3$ квадратичною функцією. Далі, обчислюється похідна (методом кінцевих різниць сусідніх точок). Якщо вона близька до нуля - ми в точці істинного максимуму. Якщо похідна велика - зрушуємося в бік її зменшення, і повторюємо ітерацію, до тих пір поки похідна не стане менше заданого порогу. Якщо в процесі ітерацій ми відходимо від початкової точки занадто далеко, то це вважається помилковим максимумом, і точка більше не вважається особливою.

Для інваріантності обчислення дескрипторів особливої точки, які будуть розглянуті нижче, потрібно визначити переважаючу орієнтацію перепадів яскравості в особливій точці. Це поняття близьке до поняття градієнта, але SURF використовує трохи інший алгоритм знаходження вектора орієнтації.

Спочатку, обчислюються точкові градієнти в пікселях, сусідніх з особливою точкою. Для розгляду беруться пікселі в окружності радіуса $6s$ навколо особливої точки. Де s - масштаб особливої точки. Для першої октави беруться точки з околиці радіусом 12.

Для обчислення градієнта, використовується фільтр Хаара. Розмір фільтра береться рівним $4s$, де s - масштаб особливої точки. Вид фільтрів Хаара показаний на рисунку 1.27.

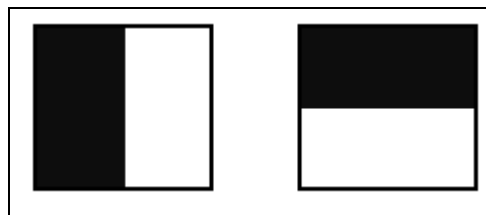


Рисунок 1.27 – Фільтри Хаара. Чорні області мають значення -1, білі +1

Фільтри Хаара дають точкове значення перепаду яскравості по осі X і Y відповідно. Оскільки фільтри Хаара мають прямокутну форму, їх значення легко розраховуються за допомогою інтегральної матриці. Для розрахунку одного фільтра довільного розміру потрібно всього 6 операцій.

Значення вейвлета Хаара dX і dY для кожної точки множаться на вагу і запам'ятовуються в масиві. Вага визначається як значення гауссіана з центром в особливій точці і сигмою рівній $2s$. Зважування на гауссіан необхідно для відсікання випадкових перешкод на далеких від особливої точки відстанях.

Далі, всі знайдені значення dX і dY , умовно наносяться у вигляді точок на площину, як показано на рисунку 1.28.

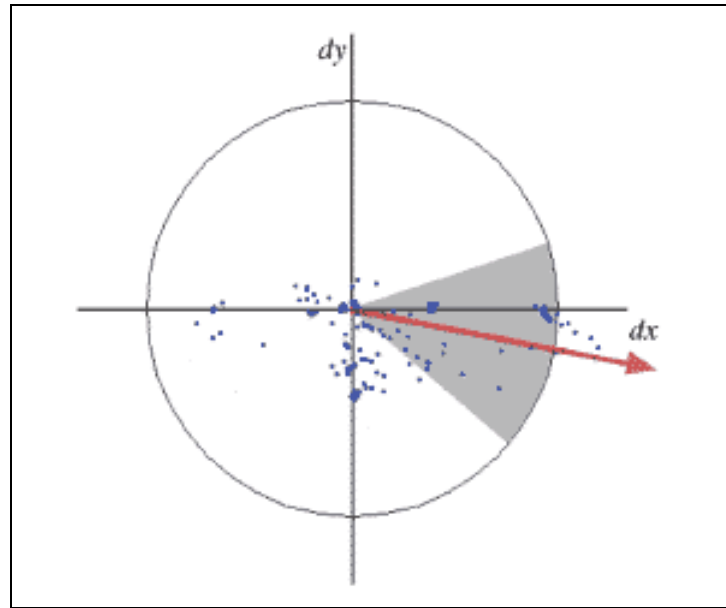


Рисунок 1.28 – знайдені градієнти у вигляді точок у просторі $dXdY$

Далі, береться кутове вікно (показано сірим на малюнку) розміром $\pi / 3$, і обертається навколо центру координат. Вибирається таке положення вікна, при якому довжина сумарного вектора для тих крапок, що потрапили у вікно – максимальна. Обчислений таким чином вектор нормується і приймається як пріоритетний напрямок в області особливої точки.

Маніпуляції з вікном потрібні для зменшення впливу шумових точок. Нижче, на рисунку 1.29 наведено приклад градієнта при ідеальному краї, і при краї з шумом.

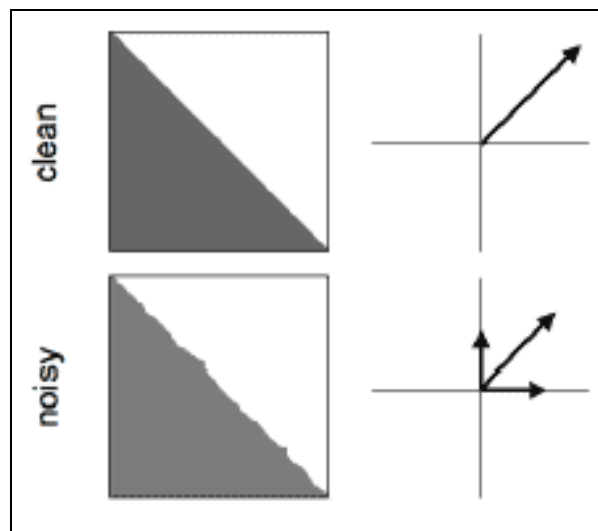


Рисунок 1.29 – Відмінності градієнтів напрямків при ідеальному краї і зашумленому

Як бачимо, шум дає додаткові градієнти в напрямках, які не збігаються з напрямом основного градієнта. Використання вікна дозволяє відсікти такі шумові точки, і більш точно обчислити істинний градієнт.

Відзначимо, що не завжди потрібна інваріантність дескрипторів щодо обертання. Метод SURF має модифікацію, в якій орієнтація особливих точок не розраховується. Така модифікація дозволяє надійно ідентифікувати точки, повернені не більше ніж на ± 15 градусів.

Дескриптор являє собою масив з 64 (в розширеній версії 128) чисел, що дозволяють ідентифікувати особливу точку. Дескриптори однієї і тієї ж особливої точки на зразку і на сцені повинні приблизно збігатися. Метод розрахунку дескриптора такий, що він не залежить від обертання і масштабу.

Для обчислення дескриптора, навколо особливої точки формується прямокутна область, яка має розмір $20s$, де s - масштаб в якому була знайдена особлива точка. Для першої октави, область має розмір 40×40 пікселів. Квадрат орієнтується уздовж пріоритетного напрямку, обчисленого для особливої точки.

Дескриптор розраховується як опис градієнта для 16 квадрантів навколо особливої точки (рис. 1.30).

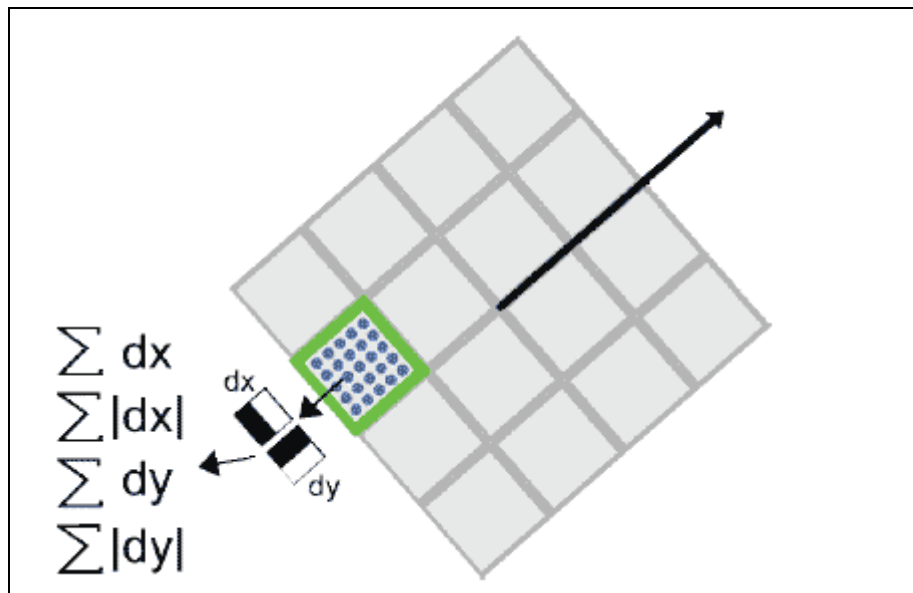


Рисунок 1.30 – Візуальний вигляд дескриптора у алгоритмі SURF

Далі, квадрат розбивається на 16 дрібніших квадрантів, як показано на малюнку. У кожному квадраті береться регулярна сітка 5×5 і для точки сітки шукається градієнт, за допомогою фільтра Хаара. Розмір фільтра Хаара береться рівним $2s$, і для першої октави становить 4×4 .

Слід зазначити, що при розрахунку фільтра Хаара, зображення не повертається, фільтр розраховується в звичайних координатах зображення. А осі отриманих координат градієнта (dX , dY) повертаються на кут, відповідний орієнтації квадрата.

Разом, для обчислення дескриптора особливої точки, потрібно обчислити 25 фільтрів Хаара, в кожному з 16 квадрантів. Разом, 400 фільтрів Хаара. Враховуючи, що на фільтр потрібно 6 операцій, виходить, що дескриптор обійдеться мінімум в 2400 операцій.

Після знаходження 25 точкових градієнта квадранта, обчислюються чотири величини, які власне і є компонентами дескриптора:

$$\Sigma dX, \Sigma |dX|, \Sigma dY, \Sigma |dY|.$$

Дві з них є просто сумарний градієнт по квадранту, а дві інших - сума модулів точкових градієнтів.

На рисунку 1.31 показано поведінку цих величин для різних ділянок зображень.

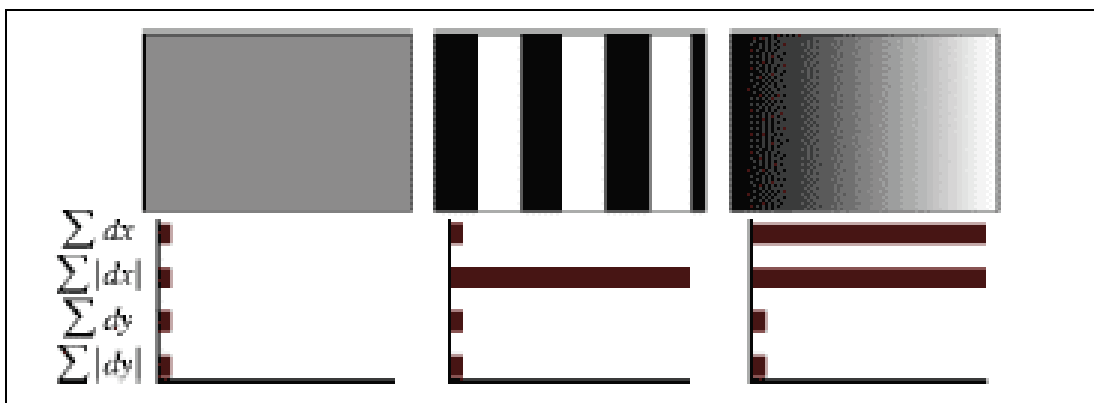


Рисунок 1.31– Значення компонентів дескрипторів для різних зображень

Рисунок 1.31 показує поведінку дескриптора для різних зображень. Для рівномірних областей - всі значення близькі до нуля. Для повторюваних вертикальних смужок - всі величини, крім другої близькі до нуля. При збільшенні яскравості в напрямку осі X, дві перші компоненти мають великі значення.

Чотири компонента на кожен квадрант, і 16 квадрантів, дають 64 компонента дескриптора для всієї області особливої точки. При занесенні в масив, значення дескрипторів зважуються на гауссіану, з центром в особливій точці і з сигмою $3.3s$. Це потрібно для більшої стійкості дескриптора до шумів у віддалених від особливої точки областях.

Плюс до дескриптора, для опису точки використовується знак сліду матриці Гессе, тобто величина $\text{sign}(D_{xx} + D_{yy})$. Для світлих точок на темному фоні, слід негативний, для

темних точок на світлому фоні – позитивний. Таким чином, SURF розрізняє світлі і темні плями (рис. 1.23).

Потрібно відзначити, незважаючи на те, що SURF використовується для пошуку об'єктів на зображенні, він сам працює не з об'єктами. SURF ніяк не виділяє об'єкт з фону. Він розглядає зображення як єдине ціле і шукає особливості цього зображення.

При цьому особливості можуть бути як всередині об'єкта, так і на фоні, а також на точках межі об'єкта і фону. У зв'язку з цим, метод погано працює для об'єктів простої форми і без яскраво вираженої текстури. Усередині таких об'єктів, метод швидше за все не знайде особливих точок. Точки будуть знайдені або на кордоні об'єкта з фоном, або взагалі тільки на фоні. А це призведе до того, що об'єкт не зможе бути розпізнаний в іншому зображенні, на іншому фоні.

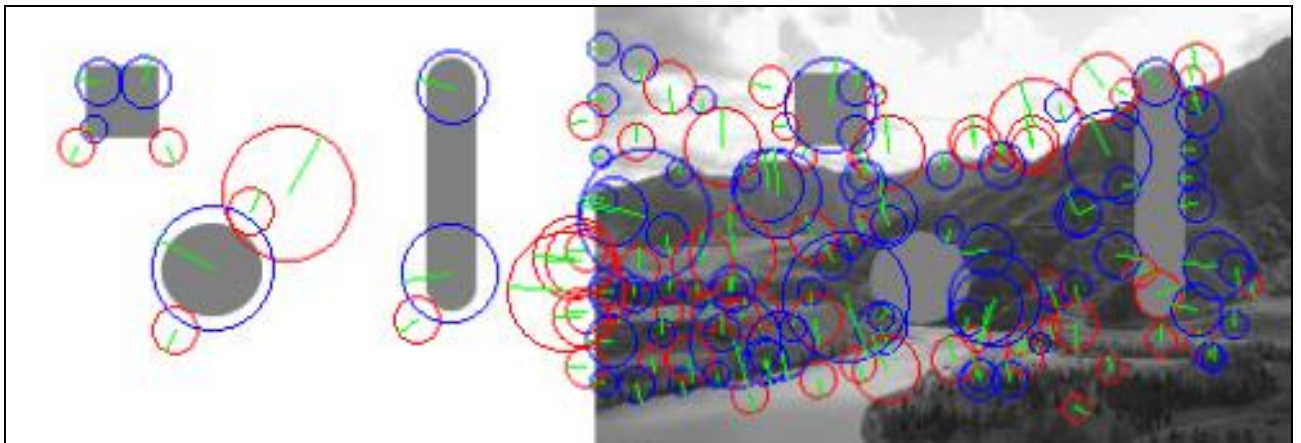


Рисунок 1.32 – прості геометричні фігури на білому фоні і на фоні природи

На рисунку 1.32 добре видно, що виявлені ознаки фігур майже ніде не збігаються. Це пов'язано з тим, що ключові точки були пов'язані з контур об'єкту і фону. Змінився фон – змінилися ключові точки.

Також до мінусів методу можна віднести те, що поріг задається абсолютним значенням, тобто на більш яскравих та контрастних зображеннях характерних ознак буде більш ніж на тусклих.

1.3 Огляд варіантів застосування

Розпізнавання плоских об'єктів може бути затребуване на рівні різних додатків побутового характеру, а також в задачах машинного зору. Застосовність можлива і в області маркетингу.

З розпізнаванням плоских об'єктів в більшості випадків тісно пов'язана технологія доповненої реальності. Це вельми нова, але вже популярна технологія має величезні перспективи і про неї варто поговорити окремо.

Якщо ви бачите речі, яких немає насправді, це означає одне з двох: ви або маєте потребу у відпочинку в маленькій затишній кімнаті з м'якими стінами, або є одним з перших користувачів революційної технології, яка в найближчий час змусить нас побачити навколишній світ зовсім по-новому. Зустрічайте доповнену реальність!

Доповнена реальність (augmented reality) - це технологія, що дозволяє накладати інформацію поверх зображення реального світу. Завдяки цьому можна обзавестися зором, як у Термінатора, а значить, не тільки швидше вистежувати ворогів, але і без зусиль знаходити точки доступу Wi-Fi.

В деякій мірі доповнена реальність - це всього лише новий спосіб отримання доступу до даних, однак вплив цієї технології на суспільство може виявитися порівнянню з ефектом від появи Інтернету. Глобальна мережа зробила інформацію більш контекстно-залежною, тобто обумовленою оточуючими умовами. Сьогодні вже не обов'язково порпатися в бібліотечних картках - пошук став швидким і зручним, а потрібні відомості можна отримати миттєво.

Доповнена реальність закріпить і багаторазово примножить ці успіхи, завдяки чому наші відносини з інформацією стануть більш ергономічними. Знайдені в мережі дані будуть автоматично доставлятися користувачам у повсякденному житті. Отже, щоб знайти хорошу кав'ярню, вам не доведеться вдаватися до допомоги пошуковиків, адже перед вашими очима миттєво опиниться список пристойних закладів, відгуки відвідувачів, а також інформація про фірмові страви та наявності Wi-Fi. Рольф Хайніч, автор книги про доповнену реальність "The End of Hardware", називає нову технологію способом створення "гіперссылки в реальному мирі" і впевнений в тому, що вона зробить життя "кінцевого користувача" більш зручним.

Якщо ви любите футбол або хокей, значить, вже напевно встигли познайомитися з доповненою реальністю. Реклама, що відображається на огорожі навколо поля, найчастіше є лише віртуальною картинкою, що накладається комп'ютером в реальному часі.

Втім, це лише квіточки, а от самі соковиті ягідки з'являться в індустрії мобільних телефонів, яка стане головною рушійною силою розвитку доповненої реальності. Новітні

телефони, оснащені GPS, камерами і компасами, навчаться визначати ваше точне положення і напрямок руху, завдяки чому інформація буде безперервно коригуватися в залежності від того, де ви знаходитесь і куди дивитесь. Вам залишиться лише дивитися на світ крізь дисплей телефону з включеною відеокамерою.

Вже сьогодні для платформ Apple iPhone, Google Android, а також версій Symbian OS і Windows Mobile існують додатки, що реалізують можливості доповненої реальності.

Найпершою програмою такого роду стала Wikitude, яка представляє собою щось схоже на Вікіпедію. Додаток доступний у версії для Google Android з кінця 2008 року, для iPhone з жовтня 2009-го і з 2010-го для Symbian OS. Програму вже скачали більше одного мільйону чоловік. Wikitude використовує інтерфейс в стилі Google Maps і дає можливість зберігати географічні координати об'єктів, що становлять інтерес (по суті, геотагінг), а потім робить цю інформацію доступною всім користувачам, які опиняться в тому ж місці. "Wikitude працює подібно просунутому навігатору, який об'єднує в собі карту, GPS і путівник", - пояснює голова відділу широкосмугового доступу стільникової компанії T-Mobile Річард Уормслі.

Щоб додаток був корисним кому-небудь, крім ентузіастів, необхідний якісний контент, а не просто аматорські фотографії, надіслані блогерами, озброївшись смартфонами. Втім, додатки, що допомагають вирішувати специфічні завдання і надають різноманітний контент, існують.

Нідерландська розробка Layar являє собою відмінний спосіб впоратися з постійно зростаючим потоком інформації, що ллється з інтернету. Layar дозволяє вибирати шари з даними, які відображаються на дисплеї телефону. Кожен шар (а зараз їх близько 300) забезпечує доступ до різних видів інформації - від порад по зніманню житла в Японії до пошуку найближчого "Макдоналдса" або станції метро в Парижі чи Лондоні. Брюс Томас з лабораторії мобільних комп'ютерів Університету Південної Австралії вважає, що "існує ще пара обмежень, які слід усунути, щоб сервіс став по-справжньому успішним". Однак, судячи з того, як швидко ранні послідовники освоїлися з доповненою реальністю, а великі мобільні бренди забезпечили підтримку для відповідних програм, ця технологія швидко набуде значну комерційну цінність.

Однак з новинкою "бавляться" не тільки виробники мобільних телефонів. Якщо у вас є комп'ютер і веб-камера, то ви можете вже зараз випробувати на собі принади доповненої реальності. У липні 2009 року читачі американського журналу Popular Science отримали незвичайний номер: якщо піднести його до веб-камери, на дисплеї ПК можна побачити тривимірне анімоване зображення вітряної електростанції, яка виступає прямо з обкладинки. А якщо ще й подути в мікрофон, то пропелери почнуть крутитися швидше!

Цей рекламний інструмент стає все більш популярним при просуванні нових фільмів та ігор. Австралійські шанувальники картини "Ніч у музеї - 2" могли піднести до веб-камери рекламний флаєр фільму і побачити тривимірною Бена Стіллера з його знаменитою дурнувато-безглуздою усмішкою - жорстоко, але ефективно. Компанія Lego застосовувала схожу технологію - упаковки з конструкторами демонстрували варіанти зібраних моделей.

У 2010 році у продажі з'явилися перші іграшки, створені з використанням доповненої реальності. Гра Sony EyePet дає гравцям можливість взаємодіяти з кумедними віртуальними звірками, використовуючи веб-камеру, а фігурки персонажів з фільму "Аватар" поставляються зі спеціальними позначками, що дозволяють бачити анімовані, літаючі, що стріляють зображення Джейка Саллі, костюма посиленою рухливості і вертольоти "Скорпіон".

Застосування доповненої реальності в таких областях, як комп'ютерні ігри, туризм, покупки, гаджети, соціальні мережі, військова справа і навіть збирання меблів, дозволить зробити рішення різних завдань більш простим і навіть веселим ділом (не в двох останніх випадках, зрозуміло). Але щоб отримати від нової технології максимум, потрібно придумати заміну стільниковим телефонам. Необхідний новий пристрій, побудований на абсолютно інших принципах. В ідеалі він повинен транслювати інформацію безпосередньо в очі, перетворюючи нас на такі собі біонічні всевидючі додатки інтернету.

На щастя, рішення вже є: Бабак Парвіз, дослідник з Вашингтонського університету, створив електронні контактні лінзи з бездротовим харчуванням, які чудово вирішують означену проблему. Кролик легко витримує носіння таких лінз протягом 20 хвилин. Втім, вчений вперто не хоче відповідати на питання про те, коли ця технологія буде доступна кожному, лише натякає, що це трапиться при його житті ... Передбачуваний час появи: кінець 2020-х років.

У 2013 році група вчених під керівництвом Янг-Юнг Парка з південнокорейського Національного інституту науки і технологій Ульсан представила м'яку контактну лінзу з вкраплених світлодіодів, для якого використовували графен і срібний нанодріт. П'ятигодинне випробування показало, що живий кролик може носити лінзи із світловипромінюючим діодом без будь-яких побічних ефектів.

Після декількох невдалих спроб американська компанія Vuzix запустила у виробництво Wgar 920 AV - перший серійний віртуальний шолом, який не соромно надіти на людях (рис. 1.33). Хоча це навіть не шолом: пристрій нагадує трішки старомодні сонячні окуляри і дозволяє одночасно бачити об'єкти доповненої реальності, а також все, що відбувається навколо вас.



Рисунок 1.33 – Віртуальні окуляри Wrap 920 AV

У реальному світі для такої технології існує безліч застосувань, і найочевидніше - в іграх. Компанія Microsoft представила проект Project Natal, який дозволить створювати ігри, де управління здійснюватиметься за рахунок рухів тіла користувача. Уявіть, що центр дії гри переміщається безпосередньо в вашу кімнату: ви можете битися з ворогами в рукопашному бою, посилаючи їх у нокаут ударами кулаків, водити гоночний болід, стискаючи руками неіснуючий кермо, і навіть ганяти м'яч, штовхаючи повітря ногами.

У 2002 році в лабораторії мобільних комп'ютерів Університету Південної Австралії, була створена гра ARQuake (версія знаменитого шутера Quake), яка використовувала елементи доповненої реальності. Для початку програмісти побудували тривимірну модель університетського містечка. Потім вони забезпечили гравця прозорим віртуальним шоломом, мобільним комп'ютером, системою визначення положення камери і поліпшеним GPS-приймачем (визначаючим позицію з точністю до 50 см, а не 5 метрів, як завжди). В результаті людина отримала можливість ходити по університету і вбивати монстрів, які ховалися за кутами і під кафедрами. Шкода, що гру не можна придбати - ми б із задоволенням побігали по рідному вузу. До речі, подібні розробки застосовуються у військових проектах вже сьогодні.

Американська "Система імітації бойових умов методом комп'ютерної симуляції" (BARS) є не що інше, як доповнена реальність на службі армії. BARS використовує ту ж концепцію, що і ARQuake, але робить це набагато більш практичним способом, допомагаючи солдатам ефективніше діяти при зіткненнях з противником в складних міських умовах. Якщо, перебуваючи в небезпечному районі, ви вирішите зупинитися і подивитися на карту, тоді будете являти собою нерухому мішень. Але якщо на вас надійти віртуальний шолом з підтримкою BARS, тоді карта просто з'явиться перед вашими очима, а завдяки алгоритмам прогнозування розташування ворога (що використовують дані про його попередні переміщення) ви навіть зможете отримати підказку про те, куди краще не пхати носа. Ця система була випробувана в Іраку і ще далека від досконалості, але її використання планувалося у американських солдатів.

Ті ж технології можуть бути використані і в навчанні. Наприклад, відкриваючи підручник, ви зможете бачити не тільки ілюстрації, але і тривимірні приклади. Уявіть, наскільки більш цікавими стали б уроки історії, якби в них використовувалися елементи доповненої реальності: ніщо так не захопило б школярів, як, скажімо, анімоване зображення лева, який бореться з гладіатором в Колізеї.

Доповнена реальність стане в нагоді людям дуже багатьох професій. Пожежні, які підходять до палаючого будинку, зможуть отримувати схеми поверхів із зазначенням положення мешканців і газових труб. Лікарям не доведеться відволікатися на перегляд рентгенівських знімків: зображення буде просто накладатися на пацієнта. Крім того, лікарі зможуть буквально заглядати всередину тіла, завдяки чому значно підвищиться точність діагностування.

Не менш приємні сюрпризи чекають тих, хто любить майструвати, а також покупців різноманітних гаджетів. Інструкції та керівництва вийдуть на новий рівень інтерактивності. Підключаючи щойно куплену аудіосистему або збираючи новенький гардероб з магазину "ІКЕА", ви зможете побачити поетапну інструкцію, що висить прямо над купою проводів або горою деталей. Це заощадить вам час і нерви, збереже сімейний спокій, а також допоможе зробити світ більш тихим місцем.

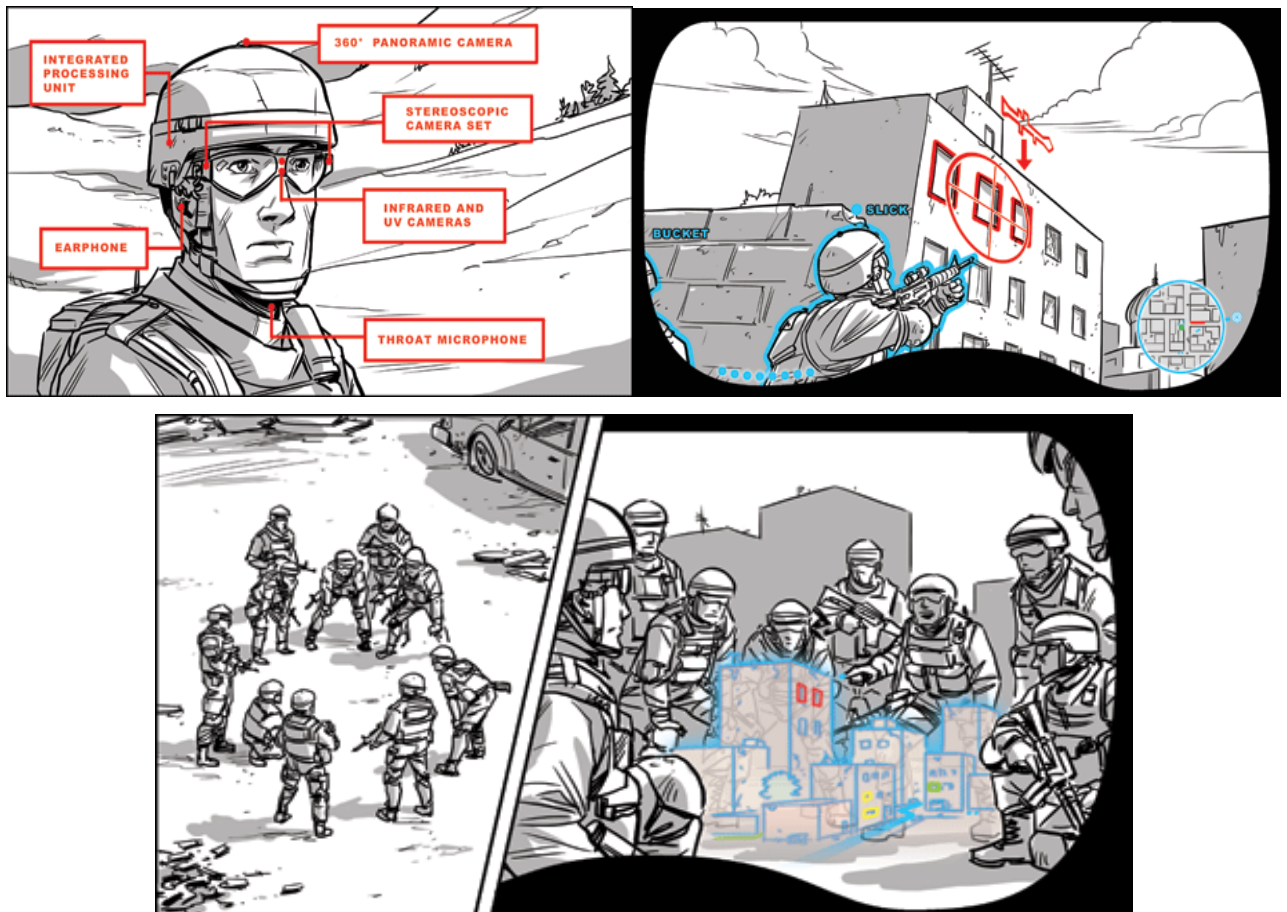


Рисунок 1.34 – Рисована схема доповненої реальності

Зрозуміло, доповнену реальність не обійшли увагою рекламодавці і маркетологи. Вони дуже хочуть "допомагати" нам здійснювати покупки, пропонуючи те, що, на їх погляд, нам хотілося б придбати. Уявіть, що, проходячи повз музичного магазину, ви будете щоразу бачити список "рекомендованих" альбомів, який підкаже, що, судячи з ваших попередніх покупок, ви терміново потребуєте альбом "Кращі хіти застільних вечірок". До речі, вже були розроблені системи, які дозволяють приміряти одяг віртуально, стоячи перед екраном у магазині.

Ден Чапмен, глава цифрового відділу Лондонського агентства BLM Red вважає: "Ми зможемо перетворити походи по магазинах у більш приємне заняття, направляючи покупців до тих речей, які їм потрібні, і демонструючи, як ці речі будуть виглядати при носінні".

Кінцева мета всього цього в тому, щоб кожна людина існував у своїй власній персоналізованій реальності і була оточена індивідуально підбраною інформацією. Подібно оточенням в соціальних мережах типу "ВКонтакте" або Twitter, приватний навколишній простір можна буде прикрашати і видозмінювати за своїм смаком. Це означає, що широке поширення доповненої реальності призведе до фундаментальних соціальних, культурних і технологічних змін.

"Найбільш цікаве повсякденне застосування доповненої реальності - це встановлення зв'язку між реальною людиною і його альтер его в соціальних мережах", - вважає Річард Лейланд, технічний аналітик, творець iPhone-додатку WorkSnug. Лейланд сподівається, що доповнена реальність відкриє доступ до інформації, що зберігається на наших сторінках у Facebook, "ВКонтакте" і т.п. Він вважає, що ми отримаємо можливість вибірково ділитися частинками свого цифрового "Я" з навколишнім світом.

Річард Уормслі з T-Mobile додає: "Спілкування в соціальних мережах веде до розмивання меж між онлайнним і офлайнним життям і в підсумку зробить можливим застосування збереженої в акантах користувачів інформації в контексті реального світу". Наприклад, ця інформація зможе з'являтися у вигляді "хмарок" з написами над головами людей. Завдяки запаморочливому успіху Twitter трансляція власних думок на весь світ стало дуже популярним заняттям, але викладати на загальний огляд (у вигляді хмарки над головою) свої плани на вечір - це, мабуть, занадто. Користувачам доведеться з обережністю ставитися до того, що вони пишуть, або обмежувати коло тих, хто має доступ до такої інформації.

"Скоро нас будуть всюди оточувати віртуальні предмети і пристрої, які замінять переважну більшість сучасних інтерфейсів", - пише Рольф Хайніч у своїй книзі "The End of Hardware". Хайніч малює світ майбутнього, де "фізичні" органи управління поступляться місцем масової персоналізації і віртуальним користувальницьким налаштуванням.

Наприклад, ви отримаєте можливість вибирати зовнішній вигляд приладової панелі автомобіля. Замість чорної пластикової панелі вашого Fiat Punto можна буде встановити елегантну приладову дошку Audi A8, і хоча ходові якості машини при цьому не зміняться, вам напевно здасться, що швидкість авто зросла.

Система SixthSense, створена дослідником Пранаві Містрі з Массачусетського технологічного інституту, - це страшне передвістя для традиційних пристроїв. Зібравши своє дітище з простих компонентів вартістю 350\$ (камера, мініпроектор, дзеркало, а також кольорові ковпачки від маркерів), Містрі може керувати інтерфейсами не гірше, ніж Том Круз у фільмі "Особлива думка", а будь-яка поверхня при цьому перетворюється в дисплей. Молодий вчений пересуває і коригує зображення, "клацає" фотографії руками і працює з віртуальними документами легко, немов вони є фізичними об'єктами. Це частково нагадує комп'ютер-стіл Microsoft Surface, але з однією важливою відмінністю - використовувати можна будь-яку поверхню.

Логічним завершенням цієї ідеї буде мир без ноутбуків, комп'ютерів, електронних книг і навіть клавіатур - досить незвичайне видовище, треба думати. Це навіть не ривок вперед, а справжній тектонічний зсув, який поки досить важко собі уявити, хоча його реалізація - лише справа часу.

Перш ніж наші голови закружляють від утопічних видінь, мабуть, варто поговорити про обмеження, властивих доповненій реальності.

"Доповнена реальність - це всього лише інтерфейс, важливість та корисність якого визначається інформацією, що проходить через нього. Без цінного і вчасно наданого вмісту, доповнена реальність буде всього лише ще одним вікном у величезній порожній кімнаті", - упевнений Адам Грінфілд, глава відділу дизайну компанії Nokia і автор книги "Everyware" про комп'ютери майбутнього. Взяти хоча б мобільний додаток Layar. "300 інформаційних шарів" - звучить вражаюче, але серед них повно таких, чия цінність, м'яко кажучи, сумнівна. Навряд чи опис другорядних пам'яток якої-небудь глибинки чи дані про висоту всіх піків в горах Швейцарії допоможуть вам знайти дорогу у великому європейському місті.

Алістер Джеффс з компанії Euerly попереджає: "Простори, в яких доводиться діяти сучасним мобільним системам доповненої реальності, настільки великі, що віддача зазвичай виявляється малоефективною і недостатньою". Співвідношення очікувань і досягнутих результатів таке, що реалізовані можливості доповненої реальності в даний час виглядають досить непоказно і навіть можуть розчарувати.

Це нагадує ситуацію з віртуальною реальністю, від якої колись чекали небачених чудес, хоча на практиці нічого особливого зроблено не було. Компанія Euerly хоче обійти цю проблему, концентруючи свої зусилля на "компактних просторах з чітко визначеними параметрами і заданим периметром" (наприклад, стадіони та парки розваг). Застосування

доповненої реальності в подібних легко керованих місцях дозволить забезпечити прийнятний рівень інтерактивності, а не розпорошувати сили, як це роблять програми на кшталт Wikitude.

Існують і інші програми з чітко сформульованими цілями. Lookator і Worksnug розраховані на фрілансерів, які воліють працювати в громадських місцях, і дозволяють знаходити Wi-Fi точки доступу. Крім цього, Worksnug вказує, де знаходяться кафе, а як бонус пропонує рейтинг якості кави і комфорту. Однак Адам Грінфілд з Nokia вважає, що в цьому криється потенційна проблема: "Доповнена реальність в її сучасному втіленні може зробити користувача уразливим. Чи думають прихильники цієї технології про безпеки, що підстерігають людину, що нерухомо стоїть, затиснувши телефон у витягнутій руці, посеред щільного міського руху?". Правда, це міркування не злякало сотні тисяч тих, хто вже скачав Wikitude і Layar.

Безсумнівно, доповнена реальність поки ще недостатньо зручна і проста, та й усю надану в ній інформацію легко знайти в Інтернеті з того ж мобільного телефону. Проте, ледве випробувавши цей спосіб взаємодії з навколишнім середовищем, ви відразу ж переконаєтеся в його перевагах порівняно зі звичайними онлайн-послугами. У міру поліпшення якості додатків і збільшення обсягу контенту, доповнена реальність перетвориться на невід'ємну частину повсякденного життя. Сьогодні Інтернет пронизує майже всі сфери нашого буття, але для користування цим благом необхідний дисплей комп'ютера або портативного пристрою. Пора зламати цю стіну між людьми та інформацією. Сьогодні мало-кому потрібен паперовий словник - набагато зручніше знайти потрібне слово в Мережі, зробивши пару кліків мишею. Точно так же втратять свою значимість і миші, коли інформація буде транслюватися людині безпосередньо в очі.

2 РОЗРОБКА АЛГОРИТМУ ІНФОРМАЦІЙНОЇ ТЕХНОЛОГІЇ

2.1 Загальний план

У якості бази для побудови алгоритму було обрано представлення об'єкта набором його характерних ознак. Тобто для шуканих об'єктів спочатку будуються моделі. Точний опис моделі і процесу її побудови буде наведений нижче.

Построєні моделі застосовуються для розпізнання об'єктів на довільному зображенні. Але для цього треба спочатку зробити відповідний опис зображення «сцени» на якій буде проводитися пошук. Опис робиться вилученням ключових ознак на «сцені» такого ж плану як і при побудові моделей об'єктів, а потім зіставлення цих ознак з шуканими.

Після отримання ключових ознак об'єкту і «сцени» робиться співвідношення одних до інших і вибір найліпших пар. Маючи пари точок на еталонному об'єкті і основному зображенні знаходяться параметри перетворень (матриця гомографії) від початкового положення об'єкта на еталонному зображенні до положення виявленого на «сцені».

Порядок дій можна описати наступним чином:

- 1) побудова моделі об'єкта;
- 2) попередня обробка зображення «сцени»;
- 3) знаходження положення об'єкта на «сцені».

При розпізнаванні об'єктів на відео потоці можливе використання інформації отриманої при розпізнаванні на попередніх кадрах, щоб уточнити положення на поточному. Тобто перед співвідношенням ключових ознак розраховується очікуване положення об'єкта базуючись на його попередніх переміщеннях.

Загальна схема розпізнавання декількох об'єктів на відео потоці показана на рисунку 2.1.

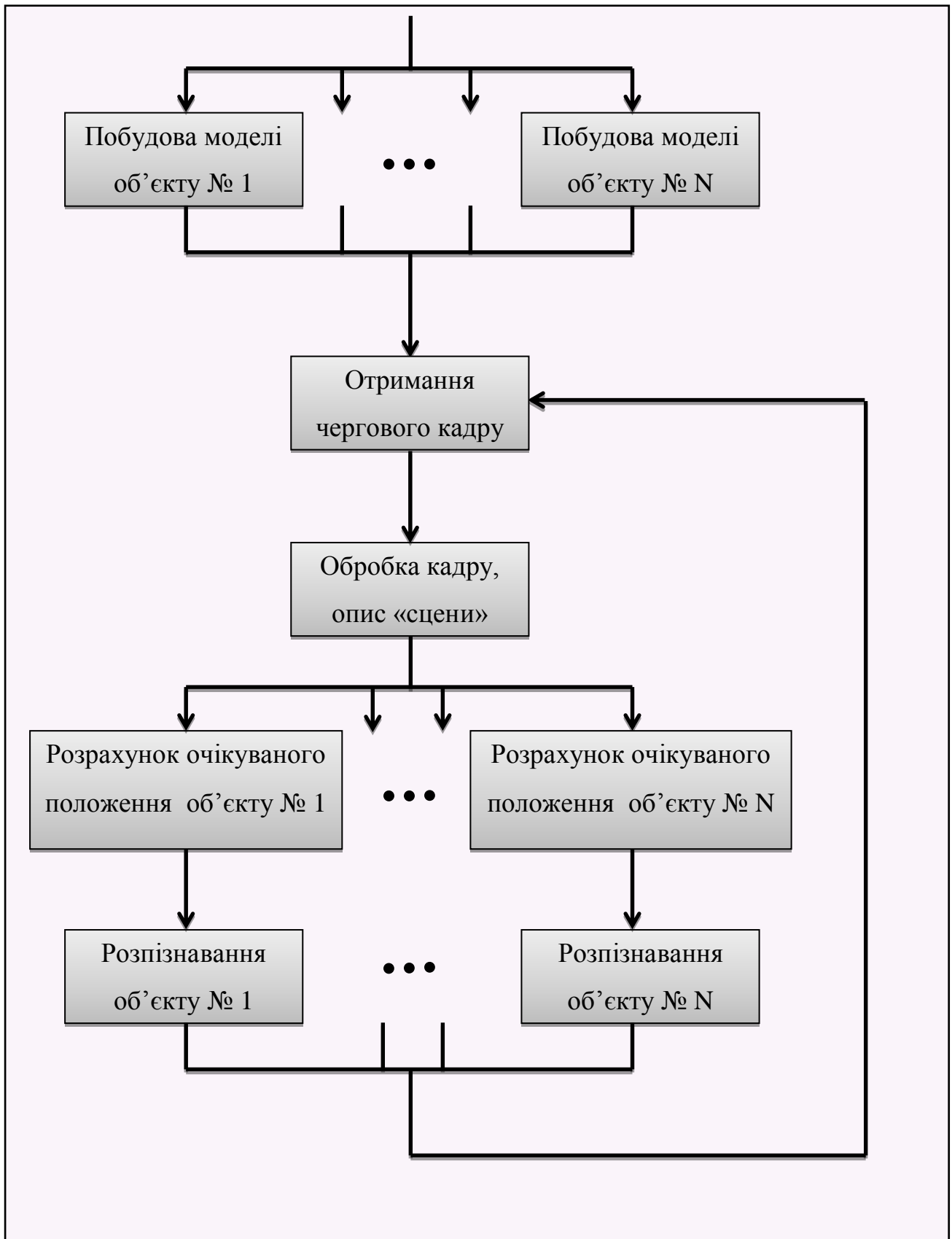


Рисунок 2.1 – Загальна схема розпізнавання об'єктів у відеопотоці

Далі розглянемо кожен етап детальніше.

2.2 Побудова моделі об'єкту

Як вже згадувалося, представлення об'єкта буде відбуватися побудовою його моделі, основою якої є набір характерних ознак.

У якості алгоритму для вибору ключових ознак було обрано метод SURF, оскільки він є одним з найбільш ефективних і швидких сучасних алгоритмів. Крім того, SURF є поширеним методом, його реалізації є в багатьох математичних бібліотеках.

Одна характерна ознака об'єкту буде включати загальні відомості про ключову точку на зображенні та її дескриптор:

- 1) положення;
- 2) масштаб (розмір фільтру при якому був знайдений екстремум);
- 3) значення гессіану (детермінант матриці Гессе у цій точці);
- 4) напрямок градієнту (напрямок найбільшої зміни яскравості);
- 5) слід матриці Гессе (чорне на білому або біле на чорному);
- 6) дескриптор (флуктуації градієнта навколо ключової точки).

Всі ці дані описані в пункті 1.2.2.

Крім характерних ознак нам ще знадобиться контур об'єкта що б після знаходження матриці гомографії накласти цей контур на зображення «сцени» помноживши на цю матрицю і таким чином отримати положення об'єкта на «сцені».

Якщо об'єкт подається зображенням на якому він повністю заповнює малюнок (а значить об'єкт прямокутний) то контуром будуть всього чотири точки: $\{0; 0\}$, $\{\text{ширина}; 0\}$, $\{\text{ширина}; \text{висота}\}$, $\{0; \text{висота}\}$.

Проте коли об'єкт не повністю заповнює зображення, чи має складну форму (відмінну від прямокутника), то він задається контуром на малюнку. Тоді доводиться не тільки запам'ятовувати контур цього об'єкту а ще й при побудові його моделі звертати увагу тільки на ті характерні ознаки які виявлені всередині контуру. А також бути обережними з тими ознаками, що близькі до краю контуру, тому що вони можуть бути залежні від фону.

І так починається побудова моделі з пошуку ключових точок використовуючи алгоритм SURF. Цей алгоритм шукає локальні екстремуми матриці Гессе тому можливе налаштування пошуку відсікаючи всі ключові точки, значення гессіану в яких менше деякого порогу. Проте цей поріг має абсолютний характер. Зображення еталону можуть бути тусклыми і переходи яскравості будуть менш виражені, тоді значення порогу треба зменшувати, щоб отримати необхідні ключові точки. А може статися і таке, що зображення еталону буде різноманітне і дуже контрастне, тоді для нього при невеликому значенні порогу

буде відібрано дуже велика кількість ключових точок що буде надлишкове і призведе до втрати швидкодії надалі.

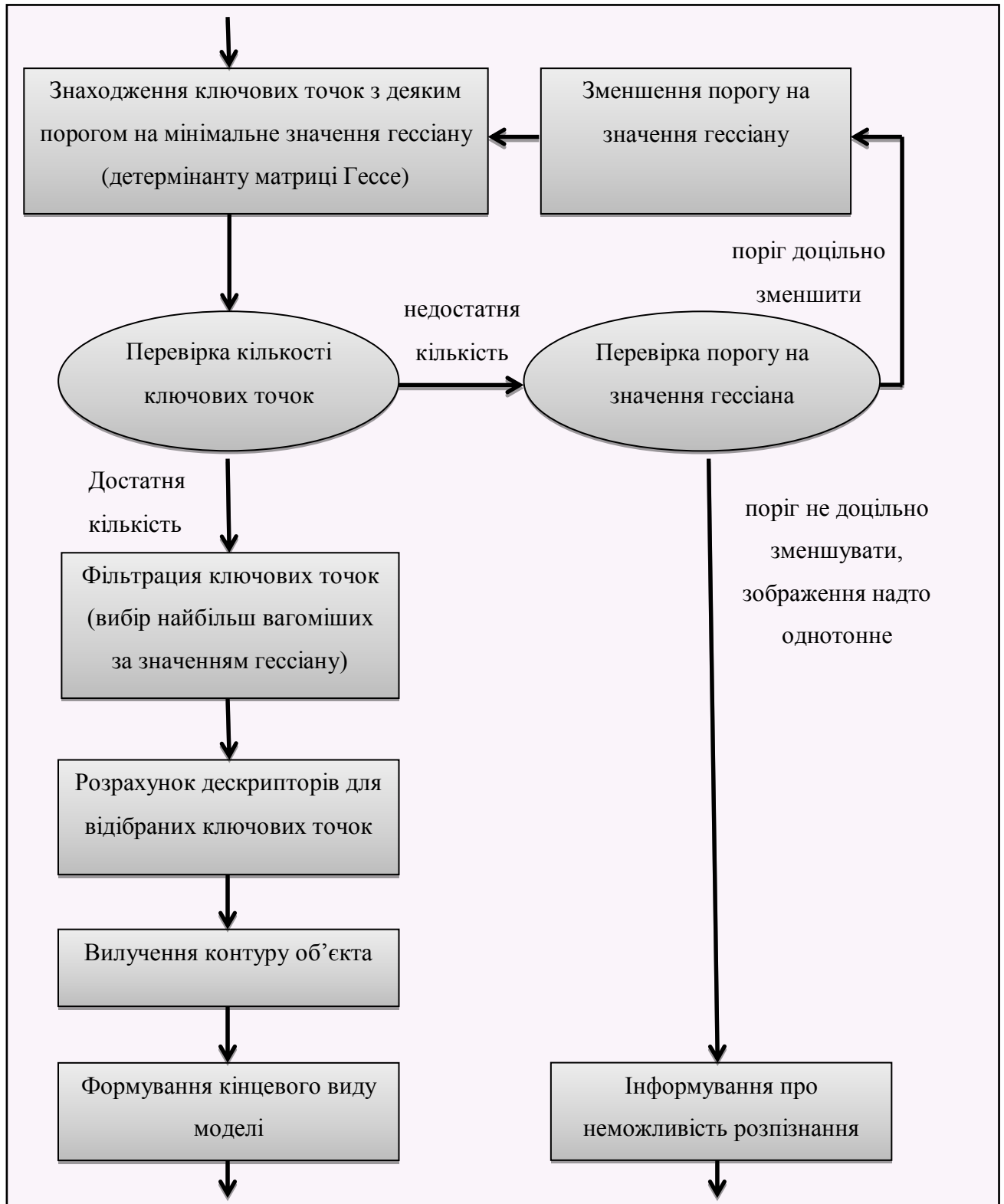


Рисунок 2.2 – схема побудови моделі об'єкта

Для подолання цієї проблеми вирішено задавати поріг досить невеликий, щоб можливо було розпізнавати і тусклі об'єкти, але після знаходження ключових точок

використовувати фільтрацію залишаючи деяку кількість найбільш вагомійших ознак. Якщо точок було знайдено менше ніж необхідна нам кількість то можна понизити поріг гессіана і знайти точки з новим порогом.

Після вибору ключових точок для кожної з них розраховується дескриптор, що являє собою флуктуації градієнта яскравості навколо точки. Процес розрахунку дескриптора описаний у пункті 1.2.2.

Значення ключових точок та їх дескрипторів групуються та зберігаються у моделі, туди ж вносяться дані про контур. У моделі об'єкту можливе використання статистики про її розпізнання на різних зображеннях. Про це ми поговоримо пізніше, а поки варто сказати що ці дані про статистику ініціалізуються початковими значеннями і після цього модель готова.

Якщо маємо декілька об'єктів то зрозуміло, що ми повинні побудувати модель для кожного з них. Побудови моделей для різних об'єктів не переплітаються, а тому при реалізації їх можна розділити на різні потоки. Це має сенс для більш швидкого старту програми але ніяк не впливає на подальшу роботу, тому що моделі об'єктів будуються тільки один раз. Надалі їх можна зберегти і загрузити. Зберігання моделей об'єктів потребує менше пам'яті ніж їх зображень та дозволяє не будувати їх знову при подальших запусках програми яка використовує цей алгоритм.

2.3 Обробка зображення «сцени»

Зображення на котрому буде проводитися розпізнавання умовно назовемо сценою. На «сцені» об'єкт може або бути або не бути (ситуація наявності декількох однакових об'єктів поки не розглядається), але заздалегідь ми не можемо знати з яким варіантом ми маємо діло. Наявність об'єкта визначається при розпізнаванні його положення, а для розпізнавання спочатку потребується отримати опис сцени у спеціальному вигляді. Про розпізнавання детальніше ми поговоримо у наступному пункті, а зараз зосередимося на попередній обробці «сцени».

Ми вже поговорили про побудову моделі об'єкту (пункт 2.1.2). У ній збережені відомості про характерні ознаки шуканого об'єкта. Завдання розпізнавання в нас вже звелось до пошуку цих ознак на «сцені». Проте де вони розміщені на сцені нам не відомо, а розраховувати значення для великої кількості можливих точок не має можливості – це досить довгий процес. Та якщо ми згадаємо що ключовими точками були локальні екстремум перепаду яскравості, то рішення стає очевидним.

Локальні екстремуми на одному зображенні об'єкта з досить великою ймовірністю залишаються екстремумами і на іншому зображенні. Тому для «сцени» знаходяться ключові точки так само як і для зображень об'єктів, а потім розраховуються їх дескриптори. Можна навіть сказати, що для сцени так само робиться модель як і для об'єкта, але тут є досить велика різниця.

Різниця моделей об'єкта та сцени полягає у тому, що при виборі ключових точок для об'єкта ми фільтрували їх і залишали певну кількість найліпших. Для сцени ми так поступити не можемо тому, що об'єкт на сцені може бути досить малий і на нього повинна припадати достатня кількість ключових точок, для розпізнавання. Мало того, об'єкт може бути тусклим і значення гессіану у точках об'єкту буде невелике, тому і поріг треба задавати досить малий. В той час саме зображення сцени може бути велике і з великою кількістю перепадів яскравості. З малим порогом на значення гессіану кількість ключових точок стає надзвичайно велика. Розрахувати дескриптори для всіх займає таку кількість часу, що про відпрацювання у реальному часу (24± кадри на секунду) можна забути.

Обхід цієї проблеми можливий наступним шляхом: візьмемо ключові точки сцени виявлені з низьким порогом (тобто їх кількість може бути дуже великою) і прорідимо їх спеціальним чином, залишивши лише необхідну кількість.

Точок з малим значенням гессіану вочевидь на порядки більше ніж точок з великим значенням. З логічних причин точки з більшим значенням гессіану є важливішими, принаймні в більшості випадків. Тому проріджування слід робити вибірково. Треба відкидати велику частину точок з малим значенням, і чим більше їх гессіан, тим більшу частину і залишати. Відкидати треба теж не аби як, а так, щоб решта точок покривала все зображення.

Таким чином маємо необхідну кількість точок з різними значеннями гессіану. Ці точки покривають все зображення і найважливіші точки залишилися.

Тепер розраховуємо дескриптори для відібраних точок. Цей набір точок з розрахованими дескрипторами і є кінцеве представлення сцени. Тут можна було б перейти до наступного розділу: «знаходження об'єкта на сцені», але є ще декілька вивертів при використанні алгоритму для відео-потoku.

Коли є прив'язка до часу і ми можемо передбачити стан об'єкта на черговому кадрі то є сенс трохи змінити схему обробки кожного кадру. По перше якщо об'єкт один і на попередньому кадрі він був знайдений, а система трекінгу ще й дає очікуване положення на поточному кадрі, то логічно з множини виявлених ключових точок відкидати ті у яких поява об'єкту мало ймовірна. Таким чином кількість точок для яких доведеться розраховувати дескриптори значно зменшується. Наступним кроком можна пришвидшити алгоритм коли об'єктів немає в кадрі. В цьому випадку слід зовсім не розраховувати дані по деяким кадрам.

Якщо поява об'єкта у кадрі буде задетектована на кадр пізніше то для більшості варіантів застосування це буде цілком прийнятно.

2.4 Знаходження положення об'єкта на «сцені»

І так ми маємо розраховані моделі для об'єкта і для сцени. Задача знайти по цим даним положення об'єкта. Задача на даному етапі зводиться до пошуку схожих характерних ознак у об'єкта і на сцені.

Ми з упевненістю можемо сказати що деякі з ознак об'єкта можуть бути не виявлені на зображенні сцени. Цьому як мінімум сприяє перекриття об'єктів та шуми. Але які саме заздалегідь ми знати не можемо, тому з початку для кожної ознаки об'єкта знаходимо найбільш відповідну на зображенні (використовуючи FLANN - fast approximate nearest neighbors), а потім з цих пар залишаємо певну кількість найліпших пар, таким чином позбавляючись від помилкових зіставлень.

Маючи вже відфільтровані пари знаходимо матрицю гомографії. Вона знаходиться по чотирьом парам точок. Використовуючи алгоритм RANSAC знаходимо загальну матрицю гомографії.

Перемножуючи отриману матрицю на кожну з точок контуру об'єкта, отримуємо розміщення контуру на сцені. І так в нас є результуюче положення шуканого об'єкта на «сцені».

Що далі? Ми ніде не робили перевірку на доцільність результату. Тобто навіть якщо об'єкта не має на сцені ми отримали його положення і нам не відомо на скільки вірна наша гіпотеза. Саме час зробити цю перевірку.

Щоб перевірити наскільки відповідають знайдені параметри перетворення (матриця гомографії) реальності візьмемо прямокутник, який охоплює контур об'єкта з початкової моделі. За допомогою матриці гомографії перетворимо цей прямокутник і отримаємо деякий чотирикутник на сцені. Якщо об'єкт був вірно знайдений, то і проекція прямокутника буде охоплювати знайдений об'єкт на сцені. Тобто фігура буде відповідати положенню прямокутника підданого «нормальним» геометричним перетворенням, перш за все вона буде випуклою.

Припустимо, що об'єкта на сцені немає, тоді результуюча матриця гомографії була отримана базуючись на випадкових точках, а значить і отримана чотирикутна фігура буде довільної форми. Якщо вона не випукла то можемо говорити про те, що об'єкта немає.

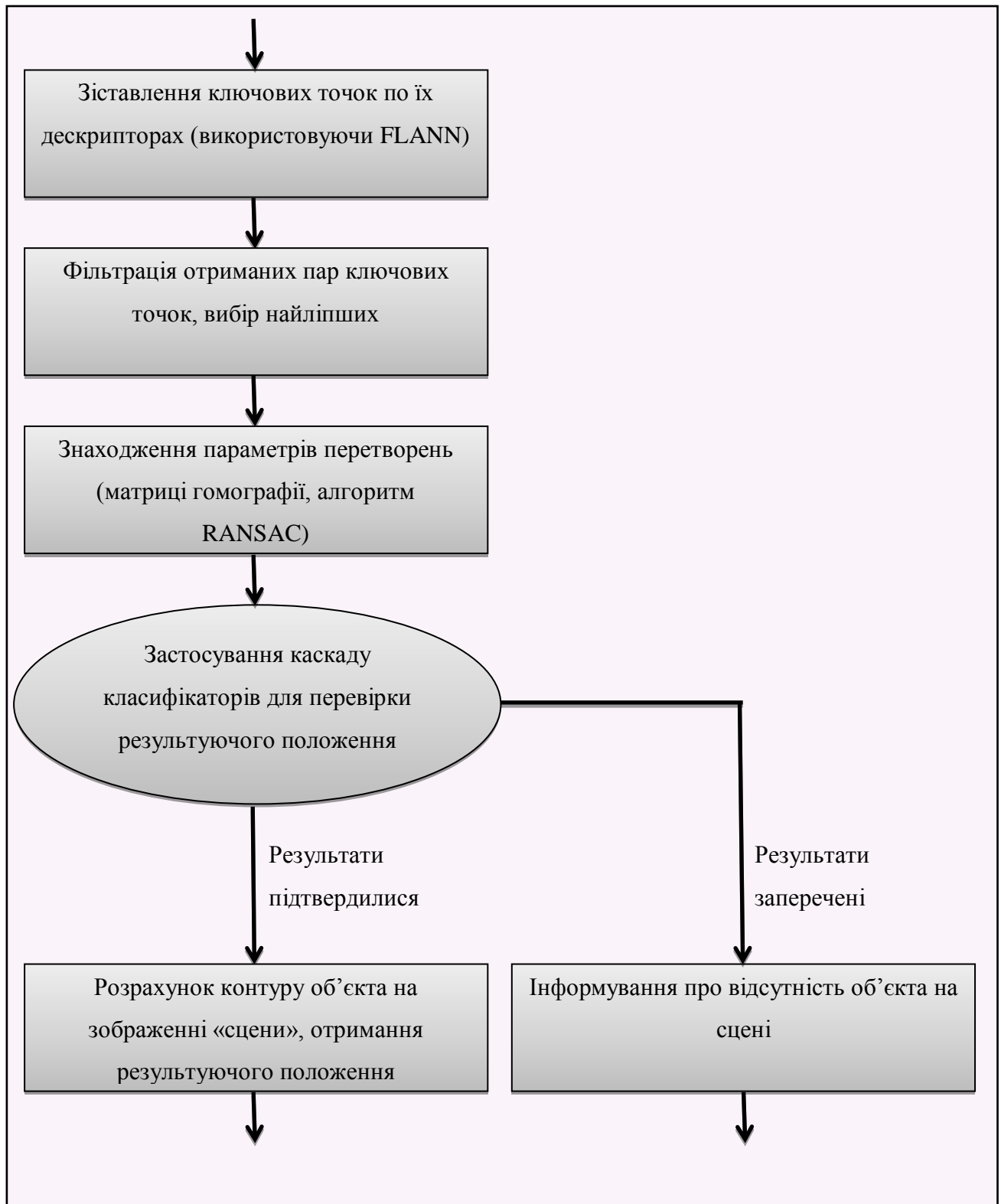


Рисунок 2.3 – Загальна схема процесу знаходження розташування об'єкта

Таким чином ми створили перший класифікатор, який перевіряє на випуклість проекцію охоплюючого прямокутника і відкидаємо помилкові знаходження. За підрахованою статистикою з 5000 різних кадрів без об'єкту було виявлено 4981 помилку. Значення близьиться до 99.62%, тобто є 0,38% що помилково знайдений об'єкт буде

прийнятий за реальний. Для деяких варіантів використань може цього буде достатньо, але наприклад при розпізнаванні з відео-потоків кожну хвилину буде виявлена наявність об'єкта у середньому 5 раз хоча його там і не буде (за 1 годину відео це близько 300 помилкових спрацьовувань).

На наступному кроці будемо працювати з результируючим положенням об'єкта на «сцені». Це положення в нас задане фігурою, яка отримана множенням матриці гомографії на контур об'єкта. Цей контур в свою чергу охоплює всі ключові точки на еталонному зображенні об'єкта. Тобто в ідеальному випадку фігура на сцені повинна так само охоплювати всі ключові точки на сцені базуючись на котрих була побудована матриця гомографії.

Якщо об'єкта не має то ця фігура буде довільного характеру і ймовірність того що всі вищеописані точки лежать всередині неї прагне до нуля при зростанні кількості точок. Проте якщо деяка незначна кількість точок знаходиться поза фігурою то це можуть бути викиди, ця ситуація нормальна і для вірного розпізнавання. Тому є деякий поріг. Експериментально цей поріг визначений у 20%. Тобто якщо у середину результируючої фігури попало менш ніж 80% точок на яких була побудована матриця гомографії то це перетворення вважається не вірним – об'єкта на сцені немає. На експериментальній статистиці отримане значення в 0.017% похибки, для 25 пар зіставлених ключових точок і порогом у 20% (мінімум 20/25 точок повинно бути у середині).

Кінцеве значення ймовірності знаходження об'єкта на сцені коли його там насправді немає можна розрахувати помноживши ймовірності класифікаторів:

$$p_{\text{загальна}} = p_1 * p_2 = 0,0038 * 0,00017 = 0,000000646.$$

Це означає що загальна ймовірність похибки дорівнює 0,0000646% (1 на 1547988). При обробці відео-потоків в середньому одна помилка буде допущена на 18 годин неперервного потоку кадрів з швидкістю 24 кадру на секунду. Такі викиди за необхідністю можна відсікати наприклад рахуючи розпізнавання правильним тільки якщо воно підтвердилося на наступному кадрі. Тоді одна похибка буде приходиться у середньому на 3210 років неперервного відео-потоків, а якщо взяти в рахунок ще ліміт на зміщення відносно першого знаходження то ця ймовірність буде надзвичайно малою.

3 ПРОГРАМНА РЕАЛІЗАЦІЯ ІНФОРМАЦІЙНОЇ ТЕХНОЛОГІЇ

3.1 Розробка демо-додатку

В процесі написання цієї дипломної був створений додаток, який демонструє роботу вищепри описаного алгоритму. Додаток реалізує алгоритм, зчитує відео-потік з камери та подає його на вхід алгоритму, обробляє результат алгоритму домальовуючи на знайденому положенні об'єкта візуальні ефекти і відображає оброблений відео-потік на дисплейну систему.

В процесі роботи додатку користувачу виводиться одно вікно з демонстрацією обробленого відео-потіку (рис. 3.1) і інше (консольне) у який відбувається вивід допоміжної та поточної інформації (рис. 3.2)).

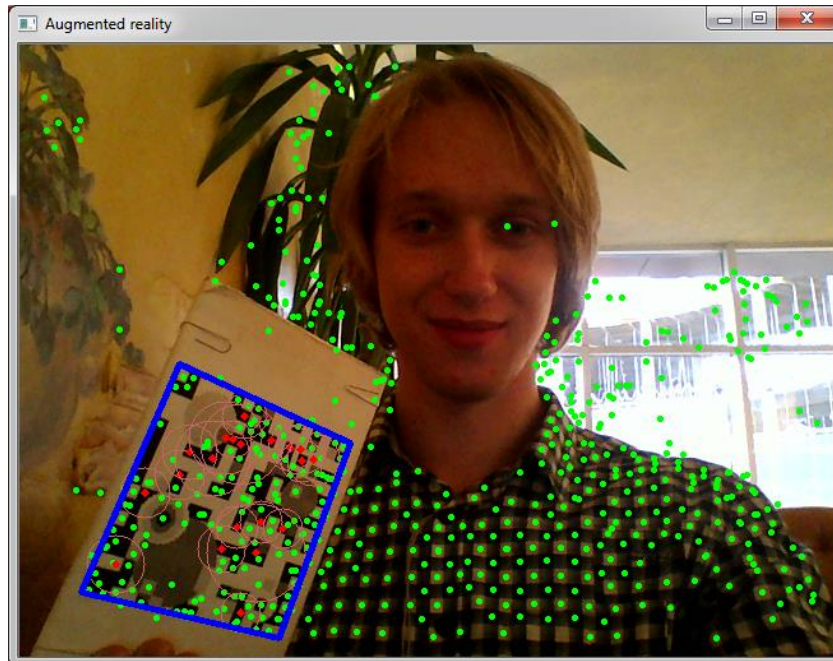


Рисунок 3.1 – Вікно відображення обробленого відео-потіку з камери

```

D:\Projects\AugmentedReality\Release\AugmentedReality.exe
.....
m - next mode
d - drawing keypoints (on/off)
x - drawing size of features (on/off)
s - stabilization (on/off)
f - filtration (on/off)
l - limitation (on/off)
b - bounding filtration (on/off)
n - noise (on/off)
- - noise +
+ - noise -
h - help
q - quit
.....
Next mode.
Imposition demo.
Next mode.
Recursion demo.
Next mode.
Static scene demo.
process: 545072 (50) | 247396
Next mode.
Dynamic scene demo.
Next mode.
Basic information demo.

```

Рисунок 3.2 – Вікно відображення допоміжної інформації

Шуканим об'єктом є заготовлене зображення, яке можна змінювати, але для показу було обрано таке зображення на якому найбільш вдало буде відпрацьовувати алгоритм (рис. 3.3). Тобто воно багате на різкі переходи яскравості та унікальні характерні ознаки.

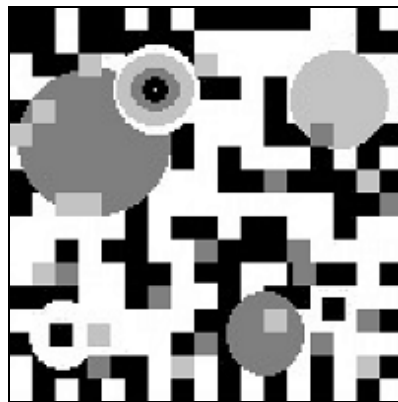


Рисунок 3.3 – Заготовлене зображення шуканого об'єкту

3.2 Посібник користувача демо додатку

В процесі роботи з додатком, крім перегляду роботи алгоритму, користувач може ще змінювати налаштування алгоритму, або змінювати візуальні ефекти. Це робиться поданням команд з клавіатури. Для більш комфортного користування кожна команда подається натиском однієї клавіші. Список команд і відповідних клавіш приводиться у консольному

вікні при запуску програми, та потім користувач може у будь-який час переглянути їх нажавши «h». Тобто «h» – це команда виводу доступних команд. Список основних команд:

- 1) h – вивід доступних команд;
- 2) m – змінити режим показу;
- 3) d – включити/виключити рисовку ключових точок;
- 4) x – включити/виключити рисовку масштабу ключових точок;
- 5) s – включити/виключити стабілізацію об'єктів;
- 6) f – включити/виключити фільтрацію знайдених пар ключових точок;
- 7) l – включити/виключити обмеження на чотирикутник, який охоплює шуканий об'єкт;
- 8) b – включити/виключити повторну фільтрація знайдених;
- 9) q – вихід з програми.

Всього зроблено п'ять режимів показів. Користувач переключає їх по кругу. Наведемо повний їх список:

- 1) відображення знайденого положення (рис. 3.4);
- 2) накладення іншого зображення замість знайденого (рис. 3.5);
- 3) рекурсивне накладення відео-потоків поверх об'єкта (рис. 3.6);
- 4) накладення статичної 3-бох мірної сцени (рис. 3.7);
- 5) накладення динамічної 3-бох мірної сцени (рис. 3.8).

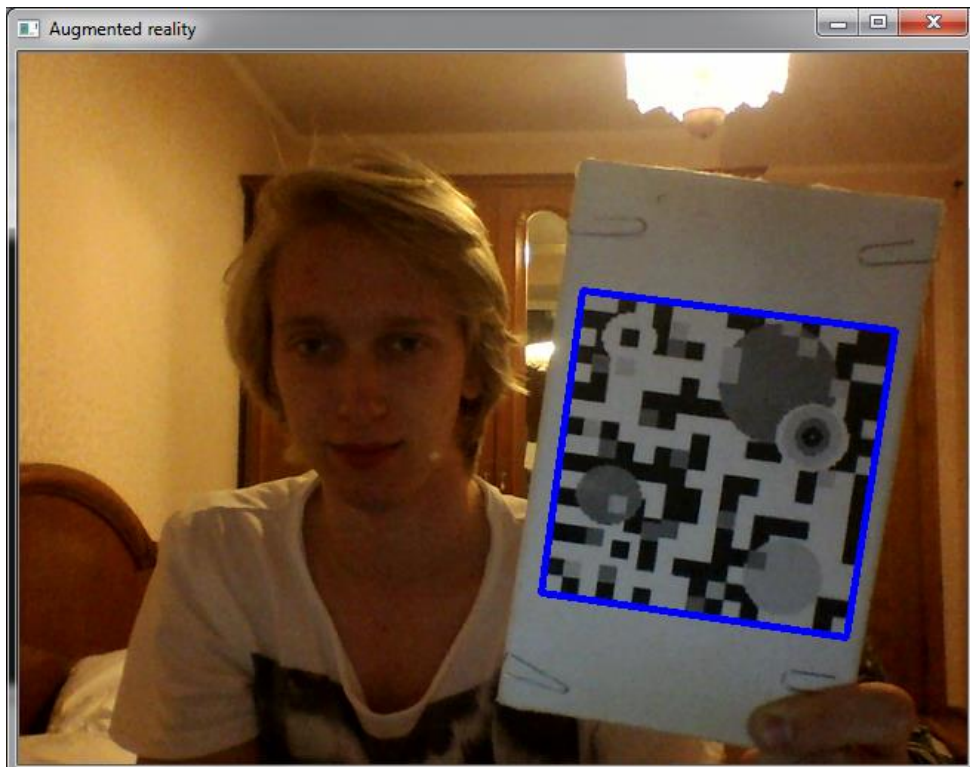


Рисунок 3.4 – Відображення знайденого положення

Режим відображення знайденого положення полягає у накладенні контуру знайденого об'єкта на відео-потік (рис. 3.4). Контур відображається синім кольором та показує положення шуканого зразка. Якщо об'єкт не знайдений то контур не відображається і відео потік виводиться без змін.

Другий режим накладення іншого зображення замість знайденого відрізняється тим, що замість контуру накладається ціле зображення (рис. 3.5). Зображення, що накладається можна змінювати натискаючи «+» та «-». Воно береться з заготовленого набору, який при бажанні можна міняти підміняючи файли з іменами «so1.jpg», «so2.jpg» і т.п.

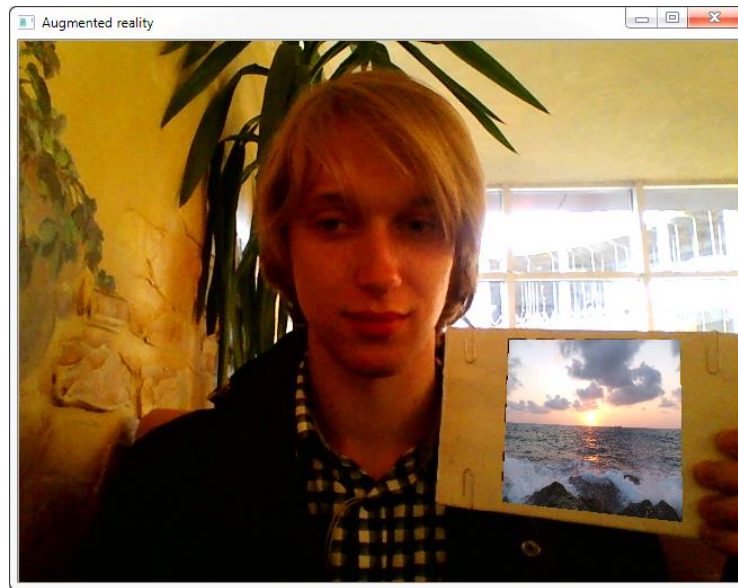


Рисунок 3.5 – Накладення іншого зображення поверх знайденого об'єкта

Наступний режим названий «рекурсивне накладення відео-потіку поверх об'єкта» і відрізняється від попереднього тим, що замість заготовленого зображення на знайдений об'єкт накладаються попередні кадри з камери створюючи таким чином рекурсію (рис. 3.6).

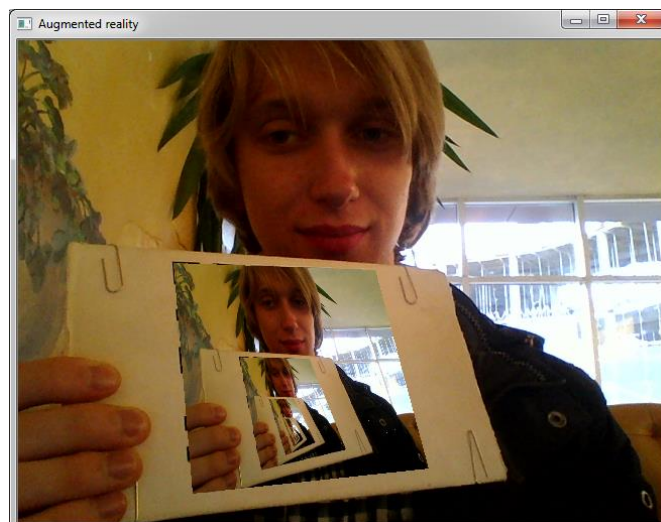


Рисунок 3.6 – Рекурсивне накладення відео-потіку

Режим накладення статичної 3-ьох мірної сцени полягає у тому що на знайдений об'єкт накладається трьохмірна модель (рис. 3.7).

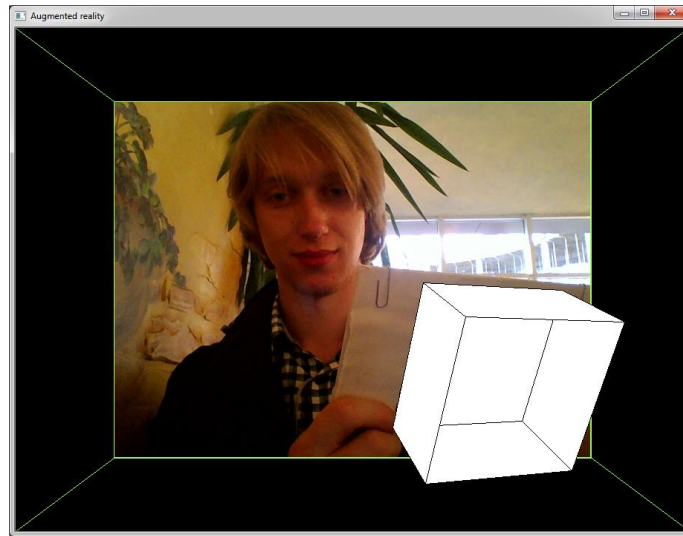


Рисунок 3.7 – Накладення статичної 3-ьох мірної сцени

Після знаходження положення у демо додатку ще обраховуються матриця моделі, представлення та проєкції. Для цих розрахунків було застосовано калібрування камери, тобто при іншій камері показ демо у даному режимі може викривитися. При цьому та наступному режимі вікно відображення розширяється створюючи додаткові поля для більшого відчуття присутності трьохмірної моделі у реальному світі.

Інший режим: «накладення динамічної 3-ьох мірної сцени» схожий на попередній але ця трьохмірна модель є рухомою (рис. 3.8).

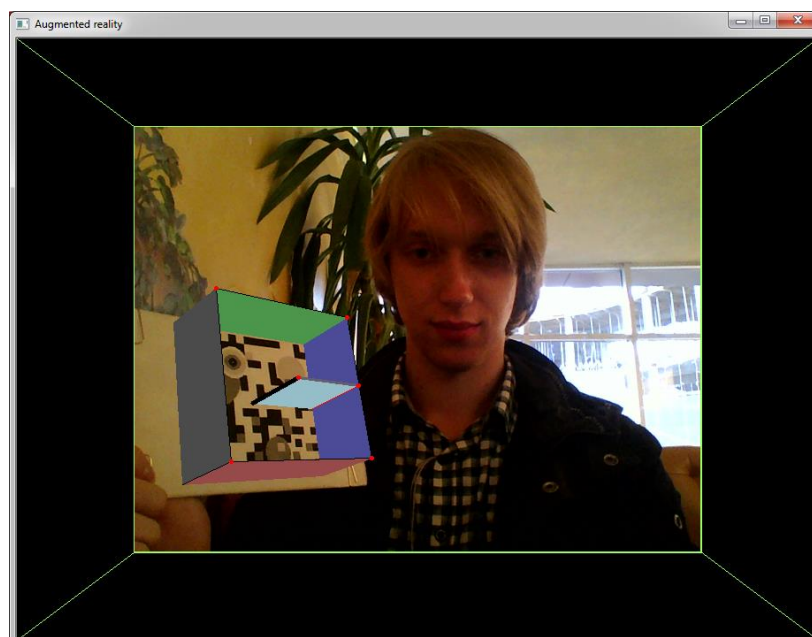


Рисунок 3.8 – Накладення динамічної 3-ьох мірної сцени

Цей режим остаточно показує можливість застосування розробленого алгоритму у доповненій реальності.

Всі режими, по наростаючому, показують різні можливості застосування описаного підходу і таким чином обґрунтовують практичну значимість роботи і застосовність у реальних ситуаціях.

3.3 Аналіз результатів

У пункті 2.4 ми вже розрахували ймовірність помилки другого роду (коли об'єкта немає на сцені, проте алгоритм видасть наявність об'єкта). Ця ймовірність дорівнює 0,0000000000000417316. Одна похибка буде приходиться у середньому на 3210 років неперервної обробки відео-потoku.

Ймовірність того що, об'єкт не буде знайдений в той час коли він знаходиться на сцені (помилка першого роду) розрахувати не вдасться, тому що це залежить від багатьох факторів. Наприклад якість знімків які подаються на вхід, погодні умови при зйомці, положення об'єкта відносно камери, тощо.

Проте можна проаналізувати допустимі повороти об'єкта при якому він знаходиться, його мінімальний розмір та допустиму наявність шуму (під шумом мається на увазі імпульсний шум – частина пікселів яка не несе інформацію про справжню картину реальності).

Спершу проаналізуємо мінімальний розмір об'єкта у кадрі при якому він стабільно знаходиться. Використовуючи демо додаток встановлене значення у 72x72 пікселі (рис.3.9), тобто на об'єкт приходиться більше 5000 пікселів.

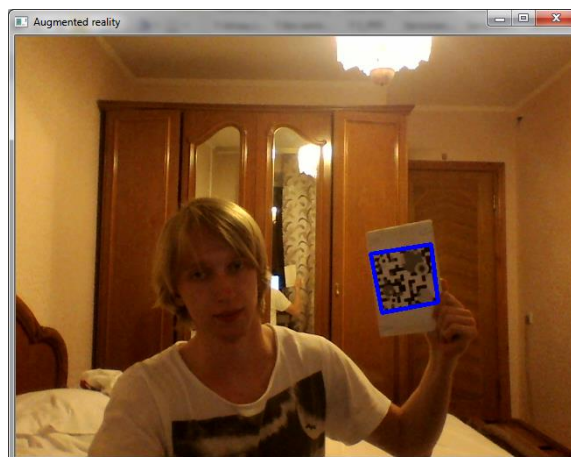


Рисунок 2.12 – Максимальна відстань на якій об'єкт стабільно знаходиться

Що до максимальних кутів при яких об'єкт стабільно знаходиться то тут можна казати про значення більше за 45° . Стабільне знаходження зберігається при поворотах близько $45-50^\circ$. Це досить гарно демонструється на рисунку 3.10.

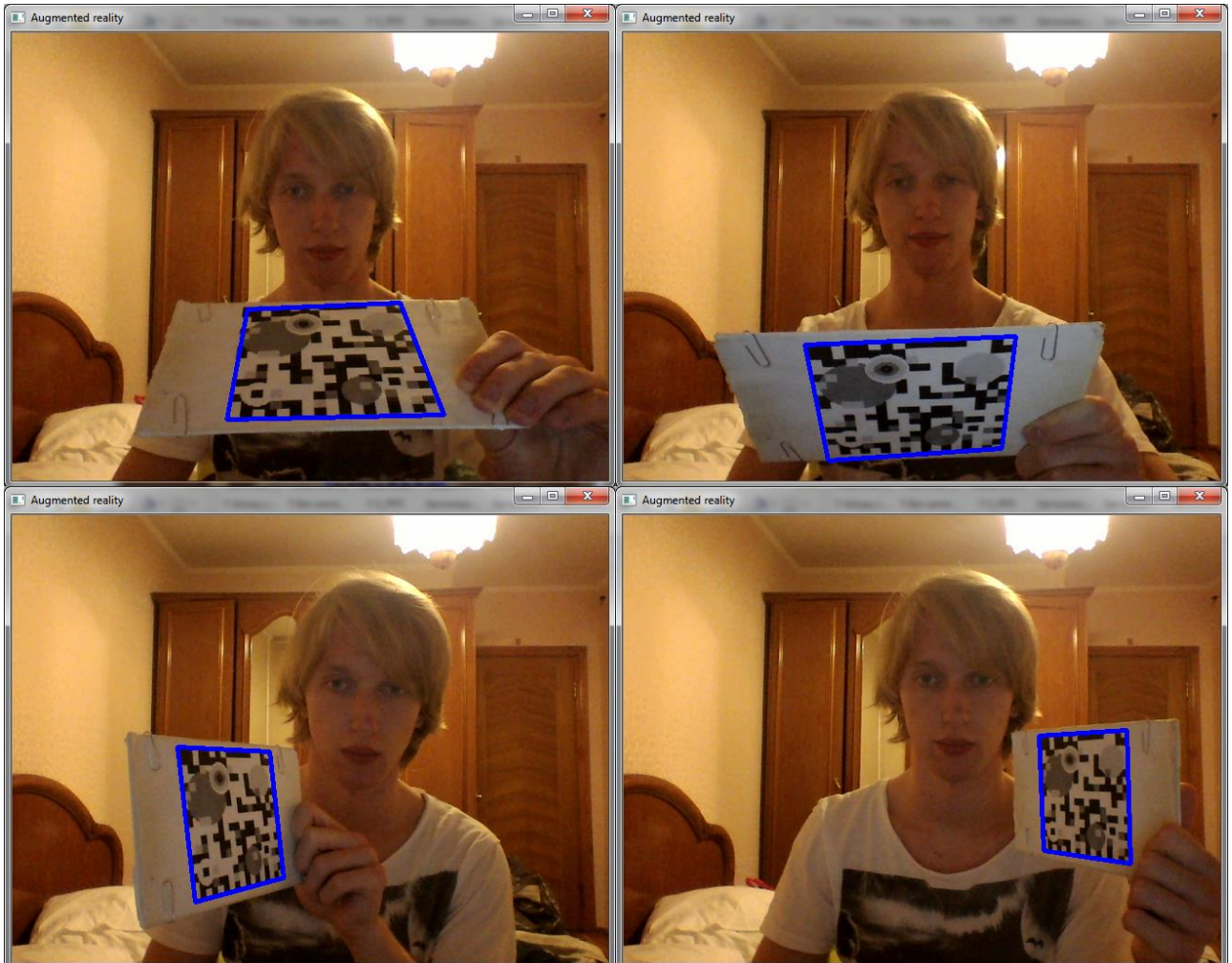


Рисунок 3.10 – Максимальні кути повороту при яких об'єкт знаходиться

Роблячи імітацію імпульсного шуму було виявлено що при 7% шуму розпізнання об'єкта змінюється незначно (рис. 3.11), зберігаються максимальні кути повороту, положення розраховується досить вдало. При 20% шуму ці параметри значно змінюються, проте об'єкт знаходиться майже завжди, коли він є на сцені і виразно виділяється (рис. 3.12). На останок можна сказати, що при 45% шуму об'єкт інколи знаходиться (рис.3.13).

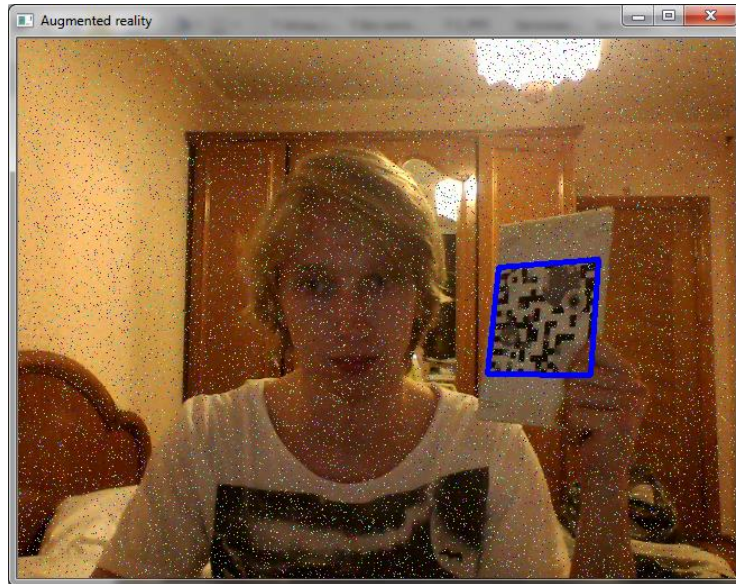


Рисунок 3.11 – Розпізнання об'єкта при 7% втрачених пікселів

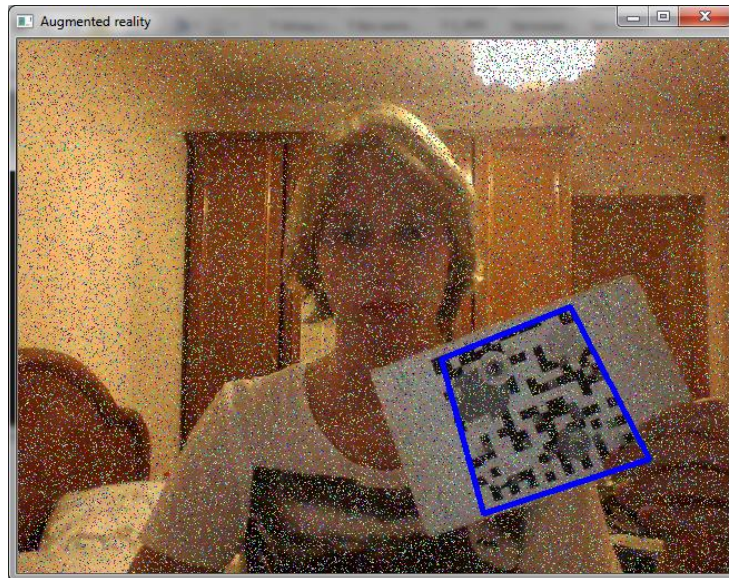


Рисунок 3.12 – Розпізнання об'єкта при 20% втрачених пікселів

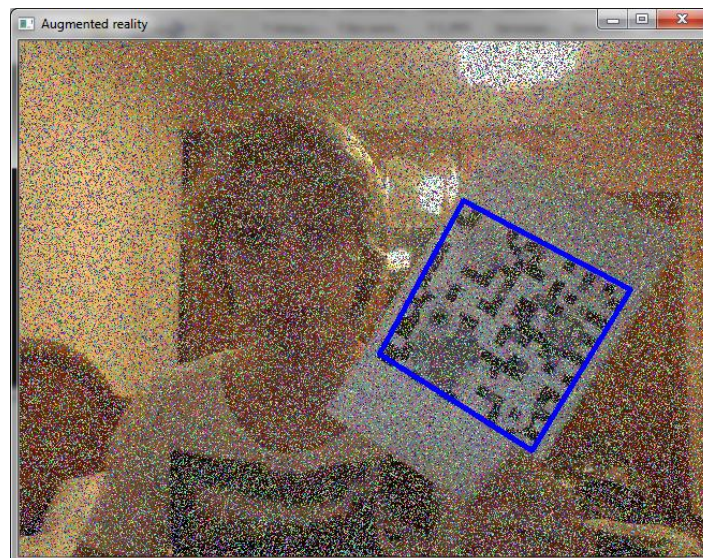


Рисунок 3.13 – Розпізнання об'єкта при 45% втрачених пікселів

Замірюючи швидкодію алгоритму був встановлений час розпізнавання одного об'єкта на одному зображенні «сцени» 640x480 пікселів. У середньому з вибірки у 1000 зображень це значення дорівнює 73,48 мілісекунд. Це означає що за одну секунду можливо обробити 13 повних кадрів. Получений час був розрахований для обробки кадру не прив'язуючись до відео-потоків та відслідковування положення об'єкта у часі. При використанні алгоритму до відео-потоків можна застосувати додаткові оптимізації щоб добитися повного ефекту роботи у реальному часі.

Всі експерименти були виконані не при ідеальних обставинах, тому ці значення можуть покращитися при використанні більш досконалої камери чи більш потужного апаратного забезпечення, або навпаки погіршитися при використанні гіршої техніки.

Основні характеристики системи на якій проводилося тестування відображені на рисунку 3.14.

Оценка:	4,6 Индекс производительности Windows
Процессор:	Intel(R) Core(TM) i7-3632QM CPU @ 2.20GHz 2.20 GHz
Установленная память (ОЗУ):	4,00 ГБ (3,84 ГБ доступно)
Тип системы:	64-разрядная операционная система

Рисунок 3.14 – Характеристики системи на якій проводилося тестування

4 ОХОРОНА ПРАЦІ ТА БЕЗПЕКА В НАДЗВИЧАЙНИХ СИТУАЦІЯХ

4.1 Аналіз потенційно небезпечних і шкідливих виробничих факторів , що впливають на персонал

Персональні ЕОМ типу IBM PC AT має наступні характеристики:

- споживана потужність 350 Вт;
- робоча напруга 220 В;
- напруга джерел живлення +12 В, -12 В, 5 В;
- робоча частота 50 Гц.

Виходячи з приведених характеристик, очевидно, що для користувача існує небезпека поразки електричним струмом у разі недбалого поводження з комп'ютером і порушення правил експлуатації (невиконання огляду відкритих частин ПЕВМ, що знаходяться під напругою або знятих для ремонту вузлів і т. д.).

Джерелами підвищеної небезпеки можуть служити наступні елементи:

- розподільний щит;
- джерела живлення;

У відповідності з [6] до легкої фізичної роботи відносяться всі види діяльності, вироблювані сидячи і не вимагаючи фізичної напруги. Робота користувача розробленого пакету програм відноситься до категорії Іа.

Згідно з [12] приміщення для ПЕОМ по ступеню небезпеки поразки людини електричним струмом відноситься до приміщень без підвищеної небезпеки (немає струмопровідної половини, вогкості, підвищеної температури, можливості одночасного дотику до корпусів устаткування з “землею” і до струмонесучих частин).

У відповідності з [7] при обслуговуванні ПЕВМ мають місце фізичні і психофізичні небезпечні, а також шкідливі виробничі чинники:

- підвищене значення напруги в електричному ланцюзі, замикання якого може відбутися через тіло людини;
- підвищений рівень статичної електрики;
- підвищений рівень електромагнітних випромінювань;
- підвищена або знижена температура повітря робочої зони;
- підвищена або знижена рухливість повітря;
- підвищена або знижена вогкість повітря;
- відсутність або недолік природного світла;
- підвищена пульсація світлового потоку;

- недостатня освітленість робочого місця;
- підвищений рівень шуму на робочому місці;
- розумове перенапруження;
- емоційні навантаження;
- монотонність праці.

Щодо до впливу на довкілля, то програмний засіб, який було розроблено під час дипломного проекту на довкілля ніяк не впливає.

Діяльність за темою магістерської роботи в процесі її виконання впливає на навколишнє природне середовище і регламентується нормами діючого законодавства: [13-18].

Основним екологічним аспектом в процесі діяльності за даними спеціальностями є процеси впливу на атмосферне повітря та процеси поводження з відходами, які утворюються, збираються, розміщуються, передаються на видалення (знешкодження), утилізацію, тощо в ІТ галузі.

Вплив на атмосферне повітря при нормальних умовах праці не оказує, бо не має в приміщенні сканерів, принтерів та інших джерел викиду забруднюючих речовин в повітря робочої зони.

В процесі діяльності виникають процеси поводження з відходами ІТ галузі. Нижче надано перелік відходів, що утворюються в процесі роботи:

- батарейки та акумулятори (малі) - III клас небезпеки;
- макулатура - IV клас небезпеки;
- матеріали пакувальні, що не вміщують целюлозу - IV клас небезпеки;
- матеріали пакувальні, що вміщують п/ет, п/пр - IV клас небезпеки;
- змінні носії інформації - IV клас небезпеки.

Наводяться вимоги зберігання виявлених за своєю роботою відходів відповідно до вимог Державних санітарних правил і норм [19].

Відходи в міру їх накопичення збирають у тару, відповідну класу небезпеки, з дотриманням правил безпеки, після чого доставляють до місця тимчасового зберігання відходів відповідно до затвердженої схеми їх розміщення. Зазначені для зберігання відходів місця чи об'єкти повинні використовуватися лише для заявлених відходів.

Не допускається зберігання відходів у невстановлених схемою місцях, а також перевищення норм тимчасового зберігання відходів.

Способи тимчасового зберігання відходів визначаються видом, агрегатним станом і класом небезпеки відходів:

- Відходи III класу небезпеки зберігаються в тарі, яка забезпечує локалізацію зберігання, дозволяє виконувати вантажно-розвантажувальні і транспортні роботи і виключає поширення в ОС шкідливих речовин;

- Відходи IV класу небезпеки можуть зберігатися відкрито на промисловому майданчику у вигляді конусоподібної купи, звідки їх автонавантажувачем перевантажують у самоскид і доставляють на місце утилізації або захоронення;

В разі тимчасового зберігання відходів у стаціонарних складах або промислових приміщеннях повинні бути забезпечені санітарно-гігієнічними етичні вимоги до повітря робочої зони згідно з [20].

Не допускається змішування відходів різних видів і класів небезпеки з будівельними і побутовими відходами, відходами дерев'яної, металевої, синтетичної тари, відходами текстильних матеріалів (старий спецодяг, ганчірки) і інш.

Проведення заготовки, здачі, переробки та реалізації металобрухту встановлені окремо [21].

Всі відходи, що утворюються в процесі діяльності/роботи, підлягають обліку.

Вимоги безпеки при поводженні з відходами:

Під час роботи з відходами (прибирання виробничих приміщень, збір і сортування, навантаження, транспортування, розвантаження та ін.) працівники та обслуговуючий персонал підприємства повинні бути забезпечені засобами індивідуального захисту та дотримуватися вимог інструкцій з охорони праці, що діють на підприємстві.

Наведено перелік деяких відходів, які передаються на утилізацію організаціям, які мають ліцензію на поводження з відходами як вторинної сировини:

- лом і кускові відходи міді, бронзи, латуні, алюмінію, свинцю;
- брухт чорних металів;
- макулатура;
- склобій;
- матеріали текстильні вторинні;
- відходи деревини кускові
- відпрацьовані фільтрувальні засоби індивідуального захисту
- відпрацьовані вогнегасники
- матеріали пакувальні вторинні

Відвантаження таких відходів здійснюється відповідно до договору (контракту).

Побутові та будівельні відходи вивозяться на полігон твердих побутових відходів міста, також відповідно до договору з комунальним дорожньо-експлуатаційним управлінням.

Особи, винні в порушенні встановленого порядку поводження з відходами (порушення правил обліку відходів, самовільне складування і видалення відходів, передача відходів в інші підприємства/організації з порушенням встановлених правил), згідно законодавства несуть дисциплінарну, адміністративну або кримінальну відповідальність.

4.2 Заходи щодо техніки безпеки

Основним небезпечним чинником при роботі з ЕОМ є небезпека поразки людини електричним струмом, яка усугубляє тим, що органи чуття людини не можуть на відстані знайти наявності електричної напруги на устаткуванні.

Проходячи через тіло людини, електричний струм надає на нього складну дію, що є сукупністю термічної (нагрів тканин і біологічних середовищ), електролітичної (розкладання крові і плазми) і біологічної (роздратування і збудження нервових волокон і інших органів тканин організму) дій.

Ступінь ураження людини електричним струмом залежить від наступних факторів:

- значення сили струму;
- електричного опору тіла людини і тривалості протікання через нього струму;
- роду і частоти струму;
- індивідуальних властивостей людини і навколишнього середовища.

Даним проектом передбачаються наступні технічні способи і засоби, застережливі поразки людини електричним струмом:

- заземлення електроустановок;
- занулення;
- захисне відключення;
- електричне розділення сітей;
- використання малої напруги;
- ізоляція струмоведучих частин;
- огорожа електроустановок.

Проведемо розрахунок заземлюючого пристрою.

Початкові дані для розрахунку заземлюючого пристрою:

- напруга установки, що заземляється, - 220В;
- режим нейтралу мережі - з ізольованою нейтралюю;
- питомий опір ґрунту – 100 Ом·м(суглинок);

- гранично допустимий опір заземлюючого пристрою - 4 Ом;
- характеристика кліматичної зони (III):
 - а) середня багаторічна низька температура, °С - від -14 до -10;
 - б) тривалість замерзання вод, дні - 150;
 - в) коефіцієнт сезонності для вертикального електроду завдовжки 3м -1,5.

Визначимо розрахунковий опір ґрунту (Ом·м) по формулі (4.1).

$$\rho_{расч} = \psi \cdot \rho = 1,5 \cdot 100 = 150 \text{ Ом} \cdot \text{м} \quad (4.1)$$

де ρ - питомий опір ґрунту;

ψ_i – кліматичний коефіцієнт, що враховує стан ґрунту під час вимірювань (таблиця 4 [7]).

Розрахуємо опір розтіканню одиночного трубчастого заземлювача по формулі (4.2).

$$R_{з.1} = \left(\frac{\rho_{расч}}{2 \cdot \pi \cdot l} \right) \cdot \ln \left(4 \cdot \frac{l}{d} \right) \quad (4.2)$$

де l – довжина заземлювача ($l=5\text{м}$);

d – діаметр труби і стрижня ($d=0,05\text{м}$);

$$R_{з.1} = \left(\frac{\rho_{расч}}{2 \cdot \pi \cdot l} \right) \cdot \ln \left(4 \cdot \frac{l}{d} \right) = \left(\frac{150}{2 \cdot 3,14 \cdot 5} \right) \cdot \ln \left(4 \cdot \frac{5}{0,05} \right) = 28,6 \text{ Ом}$$

Розрахуємо кількість паралельно сполучених одиночних заземлювачей по формулі (4.3).

$$n = \frac{R_{з.1}}{R_{доп} \cdot \eta} = \frac{28,6}{4 \cdot 0,47} = 15,2 \quad (4.3)$$

де $R_{доп}=4$. – самий допустимий опір заземлюючого пристрою;

η - коефіцієнт використання ґрунтового заземлення (для шістки заземлювачей $\eta=0,47$).

Округлятимемо отримане значення у більшу сторону $n=[15,2]=16$.

Розрахуємо довжину горизонтальної сполучної смуги по формулі (5.4).

$$L = a \cdot (n - 1) = 3 \cdot (16 - 1) = 45 \text{ м} \quad (4.4)$$

де a – відстань між вертикальними заземлювачами ($a=3\text{м}$);

n – кількість вертикальних заземлювачей ($n=16$).

Розрахуємо опір сполучної смуги по формулі (4.5).

$$R_n = \frac{\rho_{расч}}{2 \cdot \pi \cdot l} \cdot \ln\left(\frac{L^2}{d \cdot h}\right) \quad (4.5)$$

де d – еквівалентний діаметр смуги шириною $l=5$ ($d=0,05\text{м}$);

h – глибина заставляння смуги ($h=0,8\text{м}$).

$$R_n = \frac{\rho_{расч}}{2 \cdot \pi \cdot l} \cdot \ln\left(\frac{L^2}{d \cdot h}\right) = \frac{150}{2 \cdot 3,14 \cdot 5} \cdot \ln\left(\frac{45^2}{0,05 \cdot 0,8}\right) = 51,7 \text{ Ом}$$

Розрахуємо результуючий опір заземлюючого електроду з урахуванням сполучної смуги по формулі (5.6).

$$R_{ep} = \frac{R_{з.1} \cdot R_n}{R_{з.1} \cdot \eta_n + R_n \cdot n \cdot \eta_з} \leq R_{дон} \quad (4.6)$$

де η_n – коефіцієнт використання сполучної смуги (для 6-ї заземлювачей $\eta_{n=0,27}$).

$$R_{ep} = \frac{R_{з.1} \cdot R_n}{R_{з.1} \cdot \eta_n + R_n \cdot n \cdot \eta_з} = \frac{26,6 \cdot 51,7}{26,6 \cdot 0,27 + 51,7 \cdot 16 \cdot 0,47} = 3,47 \text{ Ом}$$

$3,47 < 4 \Rightarrow$ умова забезпечення електробезпеки персоналу виконується.

Таким чином, остаточна кількість заземлювачей 15 шт.

4.3 Заходи, що забезпечують виробничу санітарію і гігієну праці

Підвищення працездатності людини і збереження його здоров'я забезпечується стабільними метеорологічними умовами.

Мікроклімат виробничих приміщень – це поєднання температури, вологості і швидкості руху повітря, а також температури навколишніх поверхонь. Значне коливання параметрів мікроклімату приводить до порушення систем кровообігу, нервової і пітovidільної, що може викликати підвищення або пониження температури тіла, слабкість, запаморочення і навіть непритомність.

В приміщенні для виконання робіт операторського типу, пов'язаних з нервово-емоційною напругою, проектом передбачається дотримання наступних нормованих величин параметрів мікроклімату (див. таблицю 4.1).

Таблиця 4.1 - Оптимальні параметри мікроклімату в робочій зоні виробничого приміщення для категорії робіт 1

Період року	Температура, оС	Відносна вологість %	Швидкість руху повітря, м/с
Холодний	22.24	40.60	0,1
Теплий	23.25	40.60	0,1

Оскільки в приміщенні немає джерел виділення шкідливих речовин, можна використовувати природну вентиляцію. Площа приміщення складає 32 м². Для забезпечення прийнятних параметрів мікроклімату в приміщенні з такою площею можна використовувати 1 кондиціонер типу БК-2000.

Спектр випромінювання монітора комп'ютера включає рентгенівську, ультрафіолетову, інфрачервону області, а також широкий діапазон хвиль інших частот. Небезпека рентгенівського проміння нехтує мала, оскільки цей вид випромінювання поглинається речовиною екрану.

Для зниження дії електромагнітного випромінювання пропонується захист часом і відстанню. Захист часом передбачає обмеження часу перебування людини в зоні дії полів. Тривалість роботи на ПЕОМ повинна складати не більше 3.5–4.5 години.

Також необхідно забезпечити раціональне освітлення в робочому приміщенні. В проекті, що розробляється, передбачається використовувати суміщене освітлення. В світлий

час доби приміщення освітлюватиметься через віконні отвори, в решту часу використовуватиметься штучне освітлення.

Штучне освітлення в робочому приміщенні передбачається здійснювати з використанням люмінесцентних джерел світла в світильниках загального освітлення, оскільки люмінесцентні лампи володіють високою світловою віддачею до 75 Лам/Вт і більш, тривалим терміном служби до 10000 годин, спектральним складом випромінюваного світла, близьким до сонячного.

Зорова робота оператора ПЕВМ відповідно до [10] відноситься до розряду Va з світловим потоком $\Phi_{л}=3120$ кожна. Нормована освітленість на робочому місці (E_n) при загальному освітленні складає 200 лк.

Проведемо розрахунок кількості світильників в робочому приміщенні завдовжки $a=6$ м, шириною $b=3$ м, заввишки $c=4$ м. Формула розрахунку штучного освітлення при горизонтальній робочій поверхні методом світлового потоку (4.7):

$$\Phi_{л} = \frac{E_n \cdot S \cdot Z \cdot K}{N \cdot U \cdot M} \quad (4.7)$$

де $\Phi_{л}$ – світловий потік, Лм;

E_n – нормована освітленість;

S – площа підлоги, кв.м;

$Z=1.1-1.3$ - поправочний коефіцієнт світильника (для стандартних світильників);

K – коефіцієнт запасу, що враховує зниження освітленості в процесі експлуатації світильників;

N – число світильників;

$U=0.55-0.6$ – коефіцієнт використання, залежний від типу світильника, показника індексу приміщення і др.;

M – число ламп в світильнику.

З формули (4.7) виразимо N і визначимо кількість світильників для даного приміщення:

$$N = \frac{200 \cdot 18 \cdot 1,2 \cdot 1,5}{3120 \cdot 0,6 \cdot 2} = 1,7$$

Виходячи з цього, рекомендується використовувати 2 світильники. Світильники слід розміщувати рядами, бажано паралельно стіні з вікнами. Схема розташування світильників зображена на рис. (4.1).

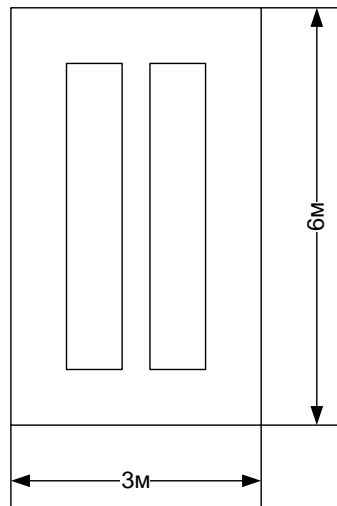


Рисунок 4.1 – Схема розташування світильників

Зниження шуму можна добитися раціонально розпланувавши приміщення, установкою устаткування на спеціальні амортизуючі прокладки. Згідно вимогам [9] рівні звуку не повинні перевищувати 50 дБ.

Для зниження стомлюваності обслуговуючого персоналу в приміщеннях, де розташовані обчислювальні засоби передбачаються використовувати спокійні колірні поєднання і покриття, що не дають відблисків. Від електромагнітного випромінювання, витікаючого від ПЕОМ, використовуються захисні екрани.

Для забезпечення чистоти повітря і відповідних мікрокліматичних умов пропонується застосувати приточування-витяжну вентиляцію. Для зменшення дії шкідливих речовин і загазованості для роботи з розплавленими матеріалами робоче місце забезпечується примусовою витяжною вентиляцією. Цей метод забезпечує притоку потрібної кількості свіжого повітря ($30 \text{ м}^3 / \text{ч}$ на одного працюючого).

Кількість повітря, яка необхідна подавати в приміщення для забезпечення необхідних параметрів повітряного середовища, визначається на підставі кількості тепла, вологі і шкідливих речовин, що поступають в приміщення, а також враховуючи видалення повітря місцевими відсмоктуваннями від устаткування, загальнообмінною вентиляцією.

4.4 Рекомендації по пожежній профілактиці

Пожежі представляють небезпеку для життя людини і зв'язані як з матеріальними втратами, так і з відмовою засобів обчислювальної техніки, що спричиняє за собою порушення ходу технологічного процесу.

Горючими матеріалами в приміщенні, де розташовані ПЕОМ, є:

- поліамід - матеріал корпусу мікросхеми. Горюча речовина. Температура samozapalennya 420 °C, енергія запалення 2мДж;
- полівінілхлорид - ізоляційний матеріал. Горюча речовина. Температура samozapalennya 480 °C, енергія запалення 50мДж;
- склостоліт ДЦ - матеріал друкарської платні. Складногорючий матеріал;
- пластикат кабельний No.489 - матеріал ізоляції кабелю. Складногорючий матеріал. Температура samozapalennya 1500 °C;
- плита деревостружкова - будівельний і обробний матеріал, матеріал з якого виготовлені меблі. Складнозапалений матеріал. Показник горючості 1.8;
- папір – довідкова і робоча документація, література. Горючий матеріал. Показник горючості більше 2.1.

Відповідно до [11] приміщення відноситься до категорії В (пожежовибухонебезпечної).

Джерелами запалення можуть бути:

- іскри при замиканні і розмиканні ланцюгів;
- іскри і дуги коротких замикань;
- перегрів від тривалого перевантаження і наявності перехідного опору;
- розряди статичної електрики.

Для того, щоб зупинити реакцію горіння, порушують умови її виникнення і підтримки. Звичайно для гасіння використовуються порушення двох основних умов сталого стану – пониження температури і режим руху газів. Пониження температури може бути досягнутий шляхом введення речовин, які поглинають багато тепла в результаті випаровування і дисоціації (наприклад, вода, порошки).

При повному тому, що згоряє органічних сполук утворюються С, SO, Н Про, N, а при тому, що згоряє неорганічних з'єднань – оксиди. Залежно від температури плавлення і тривалості реакції можуть знаходитися або у вигляді розплавів (Al O, Ti O), або підійматися в повітря у вигляді диму (P O, Na Про, MgO).

Склад продуктів неповного згоряє горючих речовин складений і різноманітний. Це можуть бути горючі речовини:

- Н, С, СН;
- атомарний водень і кисень;
- різні радикали – ОН, СН .

Продуктами неповного згоряє можуть бути також оксиди азоту, спирти, альдегіди, кетони і високотоксичні з'єднання, наприклад, синильна кислота.

Для захисту персоналу від дій небезпечних і шкідливих чинників пожежі проектом передбачено застосування промислового фільтруючого протигаза з коробкою марки В (жовтий).

До системи запобігання пожежі відносяться: запобігання утворення горючого середовища і освіти в горючому середовищі джерел запалення, забезпечення пожежебезпеки устаткування.

Щоб запобігти пожежі в обчислювальних центрах, проектом пропонується виконання наступних вимог:

- електроживлення ЕОМ має автоматичне блокування відключення електроенергії на випадок перегріву системи, що може бути результатом зупинки системи охолодження і кондиціонування;
- система вентиляції обчислювальних центрів обладнується блокуючими пристроями, що забезпечують її відключення на випадок пожежі. Система обладнується вогнеперегорджуєчими клапанами;
- застосування устаткування, що задовольняє вимогам електростатичної іскробезпеки [7];
- після закінчення роботи, перед закриттям приміщення, всі електроустановки і персональні комп'ютери відключаються від сіті електроживлення;
- в приміщеннях обчислювальних центрів забороняється:
 - влаштовувати електророзетки на основах, що згоряють;
 - використовувати синтетичні доріжки і килими;
 - користуватися побутовими електронагрівальними приладами;
 - захищувати евакуаційні виходи і проходи;
 - влаштовувати на вікнах глухі ґрати;
 - залишати без нагляду включену в електромережу апаратуру, що використовується для вимірювань і нагляду.

Для протипожежного захисту проектом пропонується обладнати приміщення площею 18 м², яке відноситься до категорії В, автоматичною протипожежною сигналізацією із

застосуванням датчиків сповіщення РІД-1 (оповіщувач димовий іонізаційний) в кількості 1 штуки і застосовується в первинних засобах пожежегасінні. Площа контрольована оповіщувачем 150 м².

Крім того, необхідно проводити навчання робочого персоналу правилам пожежної безпеки.

Розрахуємо вірогідність виникнення пожежі у виробничому приміщенні у разі запалювання транзистора:

$$Q = l \cdot T \cdot R_{кз/отк} \cdot Q_{воспл} \cdot R_{зщ} \quad (4.8)$$

де l – інтенсивність відмов пожежеопасних ЕРІ;

T – час роботи пожежеопасного ЕРІ за оцінюваний інтервал часу;

$R_{кз/отк}$ - умовна вірогідність виходу ЕРІ в стан короткого замикання при його відмові;

$Q_{воспл}$ - вірогідність запалювання ЕРІ, що знаходиться в стані короткого замикання;

$R_{зщ}$ – вірогідність відмови захисту пожежеопасного ЕРІ. Якщо захист відсутній, $R_{зщ}$ приймається рівній 1.

Вірогідність виникнення пожежі у разі запалювання транзистора:

$$Q = 1 \cdot 10^{-6} \cdot 1 \cdot 10^{-4} \cdot 0.1 \cdot 1 \cdot 10^{-4} = 1 \cdot 10^{-15}$$

Розрахована вірогідність виникнення пожежі значно менше допустимої, яка складає $1 \cdot 10^{-6}$.

В даному розділі були проаналізовані небезпечні і шкідливі виробничі чинники, що роблять вплив на персонал, розроблені заходи щодо техніки безпеки, заходу, забезпечуючи виробничу санітарію і гігієну праці, а також заходи щодо пожежної профілактики.

ВИСНОВКИ

У ході виконання даної роботи було проаналізовано великий пласт знань з машинного навчання з упором у комп'ютерний зір. З неописаних у записці і не застосованих у даній роботі підходів було розглянуто такі як дерева рішень, рандомізовані ліси, різні види характерних ознак зображень (HOG, SIFT), нейронні мережі (переважно згорткові). Щодо отриманих знань, які були безпосередньо пов'язані з цією роботою то це алгоритм SURF та побудова каскадів класифікаторів.

Було розроблено алгоритм розпізнавання плоских об'єктів за наданими еталонними зображеннями. Алгоритм випробуваний на практиці, його поточний варіант було описано у роботі [9].

Наукова новизна дослідження полягає у створенні вищезазначеного алгоритму. Він заснований на зв'язці наступних компонент:

- 1) Застосування методу surf для вилучення характерних ознак зображення;
- 2) Співставлення характерних ознак з застосуванням FLANN алгоритму;
- 3) Створення власного алгоритму вибору певної кількості найліпших екземплярів з основного набору (застосовано 2 рази: при виборі найвагомійших характерних ознак для побудови моделі плоского об'єкта та при виборі найліпших співставлень характерних ознак для двох зображень);
- 4) Розрахунку гомографії (з застосуванням алгоритму RANSAC) для переходу від заданого положення до такого який об'єкт займає на сцені;
- 5) Побудови каскаду з трьох класифікаторів для відсічення невірних результатів;

Для демонстрації можливостей алгоритму було створено демо додаток. За допомогою цього демо додатку були проаналізовані результати та виявлені різні особливості алгоритму.

Практична значимість роботи полягає в експериментальному підтвердженні ефективності запропонованого підходу. При певному апаратному забезпеченні алгоритм може працювати у реальному часі. Досягнута незалежність від таких факторів, як зміщення та поворот в фокальній площині, а також часткова незалежність від таких факторів, як зміна масштабу та яскравості, перекриття об'єктів, наявність шумів.

При виконанні дипломної роботи були розроблені заходи та технічні засоби для забезпечення безпеки праці робітників. Застосування цих засобів дозволило знизити ймовірність отримання травм при роботі, а також покращити умови праці. Було проаналізовано стан робочих місць. Для домінуючого шкідливого виробничого чинника – недостатньої освітленості – був зроблений розрахунок природного освітлення. Визначений

клас вибухонебезпеки та пожежної небезпеки приміщення, виявлені можливі причини пожежі, розроблені організаційно-технічні заходи, спрямовані на профілактику пожежі, розроблений план евакуації.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАНЬ

- 1) Кормен, Т. Алгоритмы: построение и анализ; 3-ье изд. [Текст] / К. Томас, Л. Чарльз, Р. Рональд, Ш. Клиффорд; пер. с англ. – М.: ООО «И. Д. Вильямс», 2014. – 1328 с.
- 2) Кудряшов, А.П. Извлечение и сопоставление точечных особенностей: [Электронный ресурс] //Электронный научный журнал «Исследовано в России» / Режим доступа: <http://zhurnal.ape.relarn.ru/articles/2007/104.pdf>.
- 3) Bay H. Surf: Speeded up robust features/ H. Bay, T. Tuytelaars, L.Van Gool // European Conference on Computer Vision. – 2006. – P. 1:404–417
- 4) Lowe, D. G. Object recognition from local scale-invariant features [Text] / D. G. Lowe // International Conference on Computer Vision. – Corfu, Greece (September 1999). – P. 1150–1157.
- 5) Гороховатский, В.А. Модели локально-пространственного подобия структурных описаний визуальных объектов [Текст] / В.А. Гороховатский // Системи управління, навігації та зв'язку: зб. наук. пр.– К.: ЦНДІ навігації і управління. – 2010. – Вип. 3(15). – С. 82–85.
- 6) Программно-технический комплекс технологических защит (Конфигурация САОЗ). Программное обеспечение инженерно-диагностической станции ИДС. Архив компонент поставки и компонент сопровождения.
- 7) В.И.Ященко, В.И.Дедушно, К.Е.Герасименко, О.Н.Заремская. Современные подходы к модернизации оборудования АСУ ТП АЭС на примере системы аварийного охлаждения активной зоны реактора ВВЭР-440.- ЗАО «СНПО «Импульс», Украина.
- 8) В.С. Харченко, М.А. Ястребенецкий, В.Н. Васильченко Нормирование и оценка безопасности информационных и управляющих систем АЭС: регулирующие требования к программному обеспечению // Ядерная и радиационная безопасность, 2002, № 1. – С. 18 – 33.
- 9) Qt Reference Documentation (Commercial Edition).
- 10) Тель Ж. Введение в распределенные алгоритмы. Пер. с англ. В. А. Захарова. / Ж. Тель – М.: МЦНМО, 2009. – 616 с.
- 11) 12.1.005–88. ССБТ. Общие санитарно–гигиенические требования к воздуху рабочей зоны.
- 12) 12.0.003–74. ССБТ. Опасные и вредные производственные факторы. Классификация.
- 13) 12.1.009–76. ССБТ. Электробезопасность. Термины и определения
- 14) 12.1.003-83. ССБТ. Шум. Общие требования безопасности
- 15) ДБН В.2.5-28-2006. Природне і штучне освітлення

- 16) НАПБ Б.03.002-2007. Нормы определения категорий помещений, зданий и наружных установок по взрывопожарной и пожарной опасности
- 17) ДСТУ Б А.3.2-13:2011 Система стандартів безпеки праці. Будівництво. Електробезпе́чність. Загальні вимоги
- 18) Закон України «Про охорону навколишнього природного середовища»
- 19) Закон України «Про забезпечення санітарного та епідемічного благополуччя населення»
- 20) Закон України «Про відходи»
- 21) Закон України «Про охорону атмосферного повітря»
- 22) Закон України «Про захист населення і територій від надзвичайних ситуацій техногенного та природного характеру»
- 23) Водний кодекс України
- 24) ДСанПіН 2.2.7.029-99. Гігієнічні вимоги щодо поводження з промисловими відходами та визначення їх класу небезпеки для здоров'я населення.
- 25) ГОСТ 12.1.005-88 Система стандартов безопасности труда. Общие санитарно-гигиенические требования к воздуху рабочей зоны.
- 26) Законом України «Про металобрухт»

ДОДАТОК А. Електронні плакати

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
СХІДНОУКРАЇНСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМЕНІ ВОЛОДИМИРА
ДАЛЯ
ФАКУЛЬТЕТ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ ТА ЕЛЕКТРОНІКИ
КАФЕДРА КОМП'ЮТЕРНИХ НАУК ТА ІНЖЕНЕРІЇ

Магістерська робота

Методи та інформаційна технологія виявлення та класифікації
двомірних об'єктів

Виконав:
студент гр. КН-17зм
Березін Є.О.

Керівник:
доц. Недзельський Д.О.

Северодонецьк, 2019

Актуальність



- Освіта
- Відео ігри
- Продажі і маркетинг
- Мобільні додатки
- Гео-візуалізація
- робототехніка
- виробництво
- Системи розпізнавання дорожніх знаків

Постановка задачі



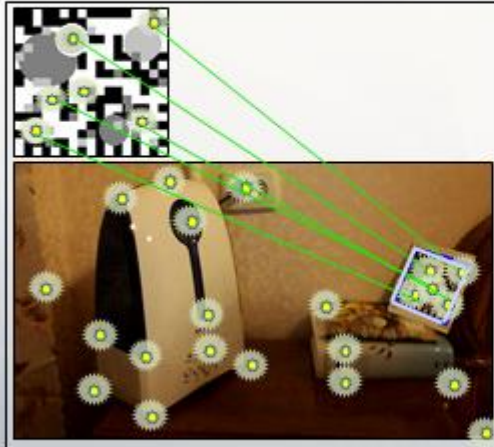
Вхід:

- Еталонні зображення об'єктів
- Відео-потік на якому буде проводитися розпізнавання

Вихід:

- Положення кожного з об'єктів в будь-який момент часу або інформування про його відсутності в області видимості

Загальна схема алгоритму



- Побудова моделі об'єктів

(В основі моделі лежить набір характерних ознак)

- Витяг необхідної інформації з поточного кадру

(Еквівалентну витяг характерних ознак на сцені)

- Пошук об'єктів на сцені

(Вибір відповідних пар ознак: сцена - еталонне зображення)

- Визначення контурів, які описують шукані об'єкти

(Розрахунок параметрів перетворення на основі отриманих пар)

- Уточнення результатів

(Позбавлення від помилкових спрацювань; стабілізація)

Побудова моделі об'єкта

$\det(H_{Harris}) = D_{xx}D_{yy} - (0.9D_{xy})^2$

- Розположення
- Значення гессіана
- Напрявлення градієнта
- Масштаб
- След матрици Гессе

$\Sigma dx; \Sigma dy; \Sigma |dx|; \Sigma |dy|$

Вхід:

- зображення
- охоплене контур

Вихід:

- набір ключових точок
- набір дескрипторів

Схема алгоритму:

- Перетворення контура
- виявлення характерних ознак
- Фільтрація по контуру (якщо необхідно)
- Вибір значимих ознак
- Розрахунок дескрипторів

Обробка зображення «сцени»

(640x480: 22000, 5000)	Без фільтрації	З фільтрацією
Виявлення	37 мс (21%)	37 мс (56%)
Фільтрація	0 мс (0%)	0.07 мс (0.01%)
Розрахунок дескрипторів	139 мс (79%)	29 мс (43%)
Загальний час	176	66

Вхід:

- Зображення
- Положення об'єктів на попередніх кадрах

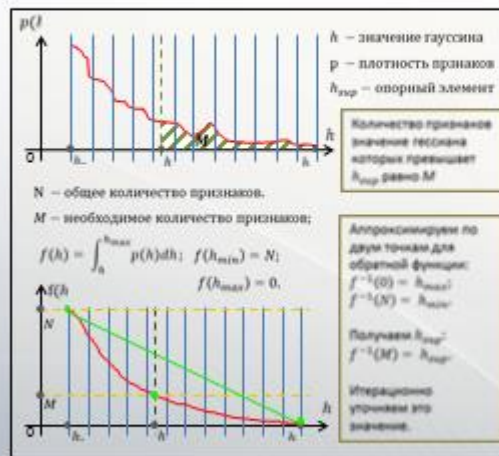
Вихід:

- набір ключових точок
- набір дескрипторів

Схема алгоритму:

- виявлення характерних ознак
- Вибір цікавлять ознак згідно бажаного покриттю
- Розрахунок дескрипторів

Проріджування характеристик



Вхід:

- Набір ключових точок
- бажане покриття

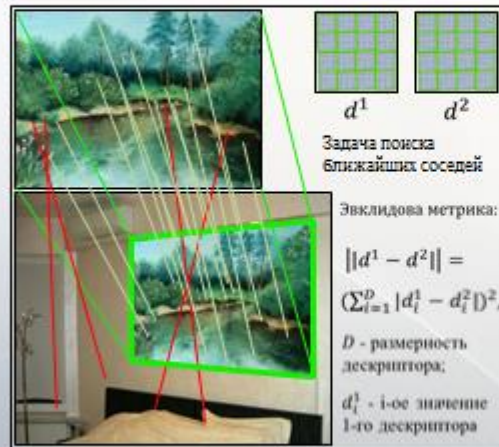
Вихід:

- Набір ключових точок згідно бажаного покриття

Схема алгоритму:

- Вибір наближеного опорного елемента
- Перестановка
- Аналіз отриманої вибірки: перехід на нову ітерацію або вихід
- (Модифікація quicksort)

Виявлення об'єкта



Вхід:

- Набір характеристик сцени
- модель об'єкта

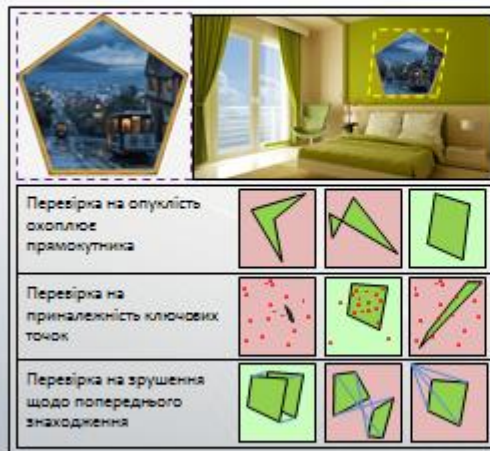
Вихід:

- Контур об'єкта на сцені

Схема алгоритму:

- зіставлення дескрипторів (Застосування FLANN)
- Вибір значимих зіставлень (Об'єкт очікується перекритим)
- Знаходження матриці гомографії (Застосування RANSAC)
- Перевірка результату (Відсікання поміжових спрацювань)

Відсікання помилкових спрацьовувань



Вхід:

- зображення сцени
- модель об'єкта
- Контур об'єкта на сцені
- попередні виявлення

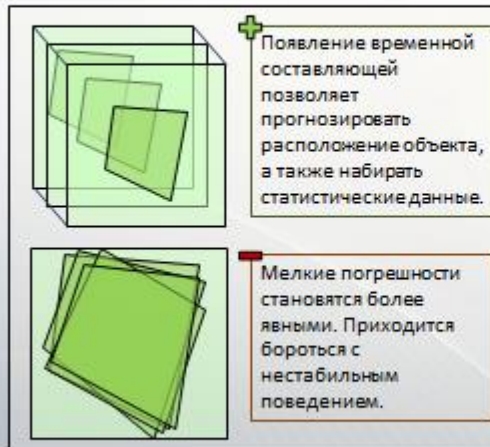
Вихід:

- Підтвердження / спростування

Схема алгоритму:

- Перевірка опуклості охоплює прямокутника
- Перевірка на приналежність ключових точок об'єкту
- Перевірка на зрушення щодо попереднього виявлення

Застосування до відео-потоків



1) Створення двох груп об'єктів (видимі і очікувані)

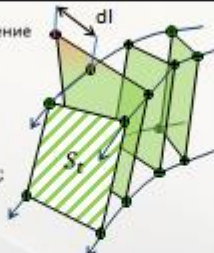
- виявлення очікуваних вноситься в окремий потік,
- Для видимих розраховується очікувана область.

2) Додаткове відсікання помилкових спрацьовувань

3) Необхідно застосування стабілізації

Стабілізація контуру

x_t^s - стабилизированное положение в момент времени t .
 l^{pro} - допустимое смещение точки контура ($0.6 \cdot S_t$).
 l^{ms0} - допустимая задержка стабилизации ($0.00006 \cdot S_t + 1,3$).
 v^s - скорость восстановления (0.1; 0.2; 0.5; 0.8; 1).



$$x_t^s = x_{t-1} + \sum_{i=t-m}^t k_i dx_i^s;$$

$$dx_i^s = \begin{cases} 0, & dl > l^{pro} + S_t \\ dx_i, & dl \leq l^{pro} + S_t \end{cases}$$

$$x_t^s = \begin{cases} x_{t-1}^s, & dx_t^s < l^{ms0} \\ x_{t-1}^s + dx_t^s + v^s, & dx_t^s \geq l^{ms0} \end{cases}$$

$$\begin{cases} \sum_{i=t-m}^t k_i = 1; \\ k_i > k_{i-1}; \end{cases}$$

Вхід:

- знайдене розташування
- Історія переміщень (Без урахування стабілізації)
- Останнє розташування (З урахуванням стабілізації)

Вихід:

- оброблене розташування
- Схема алгоритму:
- Розрахунок очікуваного контуру (З урахуванням реального)
 - Перевірка дальності між останнім від очікуваним
 - Зрушення останнього стабілізованого до очікуваного (з присвоєнням)

Аналіз результатів



Процесор: Intel(R) Core(TM) i7-3632QM CPU @ 2.20GHz 2.20 GHz
 Установлена память (ОЗУ): 4,00 ГБ (3,84 ГБ доступно)
 Тип системы: 64-разрядная операционная система

При обробці відео-потoku надходить з веб-камери отримані наступні результати:

- Оброблювані кадри мають розмір 640x480 пікселів;
- Один кадр обробляється в середньому за 63 мс;
- Максимальні кути повороту близькі до 50°;
- Мінімальний розмір розпізнається об'єкта в кадрі становить 72x72 пікселя;
- Максимально допустимий рівень імпульсного шуму становить 20%.

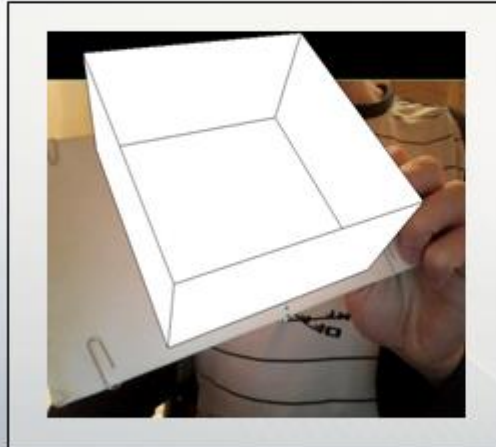
Подальший розвиток

- ✓ Розпізнавання об'єктів зі складним контуром;
(Не прямокутних)
- ✓ Застосування стабілізації руху в кадрі;
- ✓ Введення статистики за ознаками;
(Навчання в процесі)
- ✓ Застосування декількох версій еталона; (Багатогранна модель)
- ✓ Постобработка результатів методами машинного навчання;
- ✓ Випробування на інших характерних ознаках.

Підбір параметрів:

- параметри проріджування;
(Необхідне покриття, кол-во октав, характер апроксимації)
- кількість ігнорованих зіставлень;
(Можлива локалізація)
- параметри стабілізації;
(Величина історії, величина прискорення, допустиме зміщення)
- коефіцієнт очікуваної області.

Висновки



Розроблено та експериментально протестований алгоритм розпізнавання плоских об'єктів.

При певному апаратному забезпеченні алгоритм може працювати в реальному часі.

Досягнуто незалежність від зсуву і поворотів в фокальній площині, а також часткова незалежність від таких факторів, як зміна масштабу і яскравості, перекриття об'єктів і наявності шумів.

Наукова новизна полягає в певному зв'язуванні ряду компонент. Деякі з цих компонент були розроблені в процесі написання даної роботи.