

Міністерство освіти і науки України  
Харківський національний університет радіоелектроніки

Факультет комп'ютерної інженерії та управління  
(повна назва)

Кафедра електронних обчислювальних машин  
(повна назва)

**КВАЛІФІКАЦІЙНА РОБОТА**  
**Пояснювальна записка**

Рівень вищої освіти другий (магістерський)

Метод підвищення швидкодії мобільних  
відеоінформаційних систем доповненої реальності

(тема)

Виконав:

студент II курсу, групи СПМ-22-4  
Жовтоніжко Д. С.  
(прізвище, ініціали)

Спеціальність 123 «Комп'ютерна інженерія»  
(код і повна назва спеціальності)

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

Освітня програма Системне програмування  
(повна назва освітньої програми)

Керівник: проф. Можєв О.О.  
(посада, прізвище, ініціали)

Допускається до захисту

Зав. кафедри ЕОМ

(підпис)

Коваленко А.А.

(прізвище, ініціали)

2024 р.

Харківський національний університет радіоелектроніки

Факультет \_\_\_\_\_ комп'ютерної інженерії та управління \_\_\_\_\_

Кафедра \_\_\_\_\_ електронних обчислювальних машин \_\_\_\_\_

Рівень вищої освіти \_\_\_\_\_ другий (магістерський) \_\_\_\_\_

Спеціальність \_\_\_\_\_ 123 «Комп'ютерна інженерія» \_\_\_\_\_  
(код і повна назва)

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

Освітня програма \_\_\_\_\_ Системне програмування \_\_\_\_\_  
(повна назва)

ЗАТВЕРДЖУЮ:

Зав. кафедри \_\_\_\_\_  
(підпис)

“ \_\_\_\_\_ ” \_\_\_\_\_ 20\_\_ р.

## ЗАВДАННЯ

### НА КВАЛІФІКАЦІЙНУ РОБОТУ

студенту \_\_\_\_\_ Жовтоніжко Дмитру Сергійовичу \_\_\_\_\_  
(прізвище, ім'я, по батькові)

1. Тема роботи Метод підвищення швидкодії мобільних відеоінформаційних систем доповненої реальності

затверджена наказом по університету від “ 01 ” квітня 2024 р. № 257 Ст

2. Термін подання студентом роботи до екзаменаційної комісії \_\_\_\_\_ 15 червня 2024 р.

3. Вхідні дані до роботи \_\_\_\_\_  
комп'ютер з двоядерним процесором AMD Athlon II P360, тактовою частотою 2,3 ГГц  
ОС – Windows 11.

4. Перелік питань, що потрібно опрацювати у роботі \_\_\_\_\_

1) Сучасний стан мобільних відеоінформаційних систем доповненої реальності;

2) розробка методу підвищення швидкодії мобільних відеоінформаційних систем доповненої реальності;

3) дослідження методу підвищення швидкодії мобільних відеоінформаційних систем доповненої реальності;

4) висновки.

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій (слайдів) 14

---

---

---

---

---

---

---

---

---

---

6. Консультанти розділів роботи (заповнюється за наявності консультантів згідно з наказом, зазначеним у п.1 )

Найменування розділу	Консультант (посада, прізвище, ім'я, по батькові)	Позначка консультанта про виконання розділу	
		підпис	дата

### КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Термін виконання етапів роботи	Примітка
1	Аналіз літературних джерел-	02.04.24-08.04.24	
2	Вибір та обґрунтування методики дослідження	09.04.24-16.04.24	
3	Вибір інструментальних засобів	17.04.24-22.04.24	
4	Розробка моделей протоколів	23.04.24-06.05.24	
5	Проведення експериментів	07.05.24-23.05.24	
6	Оформлення матеріалів кваліфікаційної роботи	24.05.24-03.06.24	
7	Подання кваліфікаційної роботи керівникові та її попередній захист	04.06.24-07.06.24	
8	Подання кваліфікаційної роботи на рецензування	08.06.24-12.06.24	

Дата видачі завдання 01 квітня 2024 р.

Студент \_\_\_\_\_  
(підпис)

Керівник роботи \_\_\_\_\_  
(підпис)

проф. Можасєв О.О.  
(посада, прізвище, ініціали)

## РЕФЕРАТ

Пояснювальна записка кваліфікаційної роботи: 60 с., 15 рис., 6 табл., 2 дод., 13 джерел.

### ВІДЕОІНФОРМАЦІЙНІ СИСТЕМИ, ДОПОВНЕНА РЕАЛЬНІСТЬ, ОСОБЛИВІ ТОЧКИ

Метою кваліфікаційної роботи є підвищення точності та швидкодії мобільних відеоінформаційних систем доповненої реальності

У ході виконання кваліфікаційної роботи проведено аналіз детекторів спеціальних точок, що застосовуються у відеоінформаційних системах доповненої реальності. Запропоновано метод побудови дескриптора локальної особливості зображення з субпіксельною точністю, що має мінімальну довжину і не вимагає значних обчислювальних ресурсів. Експериментально встановлено доцільність запропонованого методу до різних спотворень зображень.

## ABSTRACT

Master's thesis: 60 pages, 15 figures, 6 tables, 2 appendices, 13 sources.

### VIDEO INFORMATION SYSTEMS, AUGMENTED REALITY, SPECIAL POINTS

The goal of the qualification work is to improve the accuracy and speed of mobile augmented reality video information systems

In the course of the qualification work, an analysis of detectors of special points used in video information systems of augmented reality was carried out. A method of constructing a descriptor of a local feature of an image with subpixel accuracy, which has a minimum length and does not require significant computing resources, is proposed. The expediency of the proposed method for various image distortions has been established experimentally.

## ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ .....	8
ВСТУП .....	9
1 СУЧАСНИЙ СТАН МОБІЛЬНИХ ВІДЕОІНФОРМАЦІЙНИХ СИСТЕМ ДОПОВНЕНОЇ РЕАЛЬНОСТІ.....	12
1.1 Аналіз предметної області.....	12
1.2 Створення та розвиток відеоінформаційних систем доповненої реальності.....	14
1.3 Класифікація існуючих відеоінформаційних систем доповненої реальності.....	18
2 РОЗРОБКА МЕТОДУ ПІДВИЩЕННЯ ШВИДКОДІЇ МОБІЛЬНИХ ВІДЕОІНФОРМАЦІЙНИХ СИСТЕМ ДОПОВНЕНОЇ РЕАЛЬНОСТІ.....	24
2.1 Особливості відеоінформаційних систем доповненої реальності .....	24
2.2 Визначення локальних особливостей зображення .....	25
2.3 Детектування особливих точок .....	28
2.4 Вимоги до дескрипторів локальних особливостей зображення.....	32
2.5 Побудова швидкісних дескрипторів особливих точок .....	33
3 ДОСЛІДЖЕННЯ МЕТОДУ ПІДВИЩЕННЯ ШВИДКОДІЇ МОБІЛЬНИХ ВІДЕОІНФОРМАЦІЙНИХ СИСТЕМ ДОПОВНЕНОЇ РЕАЛЬНОСТІ.....	37
3.1 Оптимізація детекторів локальних особливостей зображення .....	37
3.2 Метрики оцінки ефективності мобільних відеоінформаційних систем доповненої реальності.....	40
3.3 Результати оцінки ефективності запропонованого методу .....	42
ВИСНОВКИ .....	49
ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ .....	50

ДОДАТОК А Графічний матеріал кваліфікаційної роботи.....	52
ДОДАТОК Б Публікація .....	59

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

ВІС ДР - відеоінформаційна система доповненої реальності

ЯПД СТ – яскравий попередній детектор спеціальних точок

ЛОЗ – локальна особливість зображення

ОТЗ – особлива точка зображення

ДЛОЗ – детектор локальних особливостей зображення

МВІС ДР – мобільна відеоінформаційна система доповненої реальності

ВКФ – взаємна кореляційна функція

UML – уніфікована мова моделювання

ПЗ – програмне забезпечення

## ВСТУП

Однією із сучасних тенденцій у інформаційних технологіях є поширення технології доповненої реальності та створених на її основі відеоінформаційних систем. Доповнена реальність – це технологія, що дозволяє поєднувати об'єкти реального світу та віртуальні об'єкти, а також забезпечує їхню взаємодію. Завдання доповненої реальності – розширити взаємодію користувача з навколишнім світом, на відміну від віртуальної реальності, яка ізолює взаємодію з оточенням та переносить його у штучне середовище. В даний час технологія доповненої реальності стрімко розвивається і знаходить застосування в різних галузях діяльності людини. У медицині доповнена реальність використовується для навчання хірургів та допомоги під час проведення операцій. Наприклад, у 2013 році в Німеччині було проведено успішну операцію на печінці. Планшетний комп'ютер із встановленим спеціальним додатком транслював у реальному часі зображення печінки з накладеним інформаційним шаром тривимірних моделей проток органу, а також пухлини, яку необхідно видалити.

Доповнена реальність знаходить своє застосування у освіті. Використовуючи смартфони або планшетні комп'ютери, учні можуть побачити демонстрацію наочної операції, яку не можна виконати на аркуші паперу, наприклад, перетин тривимірної фігури площиною. Доповнена реальність є перспективним засобом для створення інтерфейсів користувача мобільних додатків, т.к. ця технологія дозволяє сприймати інформацію про положення та орієнтацію об'єктів на інтуїтивний рівень. За рахунок накладання додаткової інформації на зображення реального світу технологія доповненої реальності реалізує принципово новий тип інтерфейсу користувача, який може реагувати на навколишнє середовище. З'явився новий клас публічних виступів – інтерактивні презентації. Глядачі бачать зображення промовця на великому екрані. Якщо виступаючому необхідно

щось продемонструвати публіці, йому достатньо взяти в руки спеціальне маркерне зображення, і на нього при трансляції зображення буде виведена тривимірна модель, що демонструється, яка може реагувати на дії виступаючого. Такий підхід дозволяє заощадити виготовлення демонстрованих зразків, т.к. необхідно створювати лише тривимірні моделі, які надалі можна використовувати повторно або модифікувати для інших презентацій. Використання технології доповненої реальності для вирішення повсякденних завдань стало можливим з появою цифрових мобільних пристроїв – смартфонів та планшетних комп'ютерів, швидкодія яких дозволила забезпечити людино-машинну взаємодію в режимі реального часу.

З самого створення відеоінформаційних систем доповненої реальності ведеться боротьба за швидкодію в обробці вхідного відеоряду та розпізнавання образів. Це пов'язано з тим, що обчислювальні потужності цифрових мобільних пристроїв обмежені порівняно зі стаціонарними пристроями, оскільки висока обчислювальна навантаження призводить до підвищеного енергоспоживання і, як наслідок, скорочення часу автономного функціонування мобільного пристрою без підзарядки акумулятора.

Проте низка питань у зазначеному вище напрямку залишається невирішеною. У багатьох роботах запропоновані алгоритми та способи розпізнавання образів на зображеннях не враховують обмежені обчислювальні можливості мобільних цифрових пристроїв, і як наслідок – вимагають великих обчислювальних витрат, що при використанні мобільних цифрових пристроїв призводить до великого енергоспоживання, тепловиділення та зниження тривалості без заряду батарей.

Однією із сучасних тенденцій у інформаційних технологіях є поширення технології доповненої реальності та створених на її основі відеоінформаційних систем. Доповнена реальність – це технологія, що дозволяє поєднувати об'єкти реального світу та віртуальні об'єкти, а також забезпечує їхню взаємодію. Завдання доповненої реальності – розширити взаємодію користувача з навколишнім світом, на відміну від віртуальної

реальності, яка ізолює взаємодію з оточенням та переносить його у штучне середовище. В даний час технологія доповненої реальності стрімко розвивається і знаходить застосування в різних галузях діяльності людини. У медицині доповнена реальність використовується для навчання хірургів та допомоги під час проведення операцій. У кваліфікаційній роботі застосовані методи цифрової обробки зображень та розпізнавання образів, математичної статистики, структурного та об'єктно-орієнтованого програмування та імітаційного моделювання.

Мета роботи – підвищення точності та швидкодії мобільних відеоінформаційних систем доповненої реальності

Завдання:

1) провести аналіз детекторів спеціальних точок, що застосовуються у відеоінформаційних системах доповненої реальності;

2) запропонувати метод побудови дескриптора локальної особливості зображення з субпіксельною точністю, що має мінімальну довжину і не вимагає значних обчислювальних ресурсів;

3) експериментально встановити доцільність запропонованого методу до різних спотворень зображень.

Наукова новизна роботи полягає у запропонованому способі зменшення тривалості процесу виділення локальних особливостей зображення за допомогою детекторів кутів шляхом застосування попереднього яскравого детектора.

# 1 СУЧАСНИЙ СТАН МОБІЛЬНИХ ВІДЕОІНФОРМАЦІЙНИХ СИСТЕМ ДОПОВНЕНОЇ РЕАЛЬНОСТІ

## 1.1 Аналіз предметної області

Доповнена реальність – це технологія, що дозволяє поєднувати об'єкти реального світу та віртуальні об'єкти, а також забезпечує їхню взаємодію [1]. Ідея доповненої реальності тісно пов'язана із поняттям віртуальної реальності. Віртуальна реальність намагається створити штучний світ, який людина може відчувати та досліджувати в інтерактивному режимі, переважно через органи зору, за допомогою слуху, дотику та інших органів чуття. Доповнена реальність також створює інтерактивний світ, але для цього прагне доповнити реальний світ замість того, щоб створювати повністю штучне середовище навколо користувача. Прикладом використання технології доповненої реальності може бути траєкторія руху м'яча, що показується на екрані телевізора під час трансляції футбольних матчів [2].

Різні дослідники дотримуються різних визначень терміну «доповнена реальність» – вузкого чи ширшого. У цілому нині, наукове співтовариство на цей час дійшло згоди визначення основних частин систем доповненої реальності. У цій дисертації використовується визначення, дане одним з провідних дослідників у цій галузі Ronald Azuma: доповнена реальність – це те, що поєднує в собі реальну та згенеровану комп'ютером інформацію в реальному середовищі, інтерактивному режимі та в режимі реального часу, та поєднує віртуальні об'єкти з фізичними об'єктами. Доповнена реальність включається в ширшу концепцію змішаної реальності, яка додатково включає симуляцію, що переважно відбувається у віртуальному домені і не в реальному світі [3]. Технологія доповненої реальності стрімко розвивається і набуває все більшого поширення в різних галузях науки та техніки.

У сучасних лапароскопічних операціях зображення на ендоскопі

доповнюється зображенням, отриманим під час інтраоперативної ангіографії. Це дозволяє хірургу локалізувати пухлину всередині органа і, таким чином, мінімізувати втрати здорової тканини органа пацієнта під час операції з видалення пухлини [4].

У сучасній військовій техніці нашлемні системи індикації відображають необхідну пілоту інформацію прямо на тлі реального світу. Це дозволяє пілоту не відволікатися на панель приладів [5].

Доповнена реальність активно використовується в друкованій продукції. На сторінку друкованого видання розміщується спеціальне зображення (маркер). Читач, використовуючи мобільний пристрій (планшетний комп'ютер або смартфон), сканує зображення та отримує доступ до додаткового матеріалу, якою може виступати текст, відео, звук тощо.

Для візуалізації різних макетів або планів місцевості сучасні будівельні компанії все частіше використовують технологію, що описується. Створити тривимірну модель місцевості простіше та дешевше, ніж створити макет із картону та піску.

Технологія доповненої реальності є перспективним засобом для створення інтерфейсів користувача мобільних додатків, тому що ця технологія дозволяє сприймати інформацію про становище та орієнтацію об'єктів на інтуїтивному рівні. За рахунок накладання додаткової інформації на зображення реального світу технологія доповненої реальності реалізує принципово новий тип інтерфейсу користувача, який може реагувати на навколишнє середовище [6].

Доповнена реальність знаходить своє застосування і в освіті. Використовуючи смартфони або планшетні комп'ютери, учні можуть побачити демонстрацію наочної операції, яку не можна виконати на аркуші паперу, наприклад, перетин тривимірної фігури площиною [7].

## 1.2 Створення та розвиток відеоінформаційних систем доповненої реальності

Зародження віртуальних технологій. Виникнення віртуальних технологій зазвичай прийнято пов'язувати з Мортоном Хейлігом – голлівудським режисером, який був прихильником ідеї створення імерсійної середовища для глядача, тобто. такого середовища, яке створило б ефект якнайповнішого занурення в демонстровану реальність. Згодом технологи почали досліджувати можливість створення імерсійних середовищ і в 1955 Хейліг описав свою роботу над «кінематографом майбутнього», а в 1962 він отримав патент на винахід, названий ним «Sensorama». «Сьогодні постійно зростає попит на методи навчання та тренування людей у такий спосіб, щоб унеможливити ризики та небезпеку реальних ситуацій», — йдеться в тому патенті. У ньому описувалася віртуальна технологія, у якій візуальні образи доповнювалися рухами повітря та вібрацією [8].

Запатентований пристрій був автоматизованим робочим місцем, яке реагувало на присутність людини під час занурення в кабінку з динамічним сидінням, тривимірним зображенням і стерео-звуковим оточенням, могло додатково реагувати на зображення з супроводом запахів і вітру. Sensorama демонструвала глядачеві поїздки на мотоциклі вулицями Брукліна. Наступну спробу створення пристрою, який дозволив би занурюватися в імерсійне середовище, було здійснено у 1961 році. Це був секретний військовий проект, в ході реалізації якого вчені створили пристрій, названий Headsight, що складався з магнітної системи відстеження положення голови користувача, відеошолома з дисплеєм і трансляційних камер. Початкова мета проекту – створити апарат, який дозволив би віддалено вивчати будь-які місця, де особисто побувати з якоїсь причини було б важко чи небезпечно [9].

У 1968 році комп'ютерний спеціаліст і ад'юнкт-професор Гарварда Айван Сазерленд зі своїм студентом Бобом Спрауллом винайшли інший

пристрій, який отримав назву «Дамоклів меч». Це була перша система доповненої реальності з урахуванням головного дисплея. Вона була примітивна, а її вага була така велика, що вона підвішувалась до стелі. «Дамоклов меч» був частиною системи нічного управління гелікоптером компанії Bell Helicopter і був обладнаний датчиками повороту та нахилу голови, що давало можливість взаємодії з віртуальним інтерактивним середовищем [10]. 1981 року відбулася перша спроба мініатюризації систем доповненої реальності. Стів Манн, відомий як піонер комп'ютерів, що носилися, об'єднав шолом-камеру з видошукачем і рюкзаккомп'ютер, що носив, що дозволяло робити накладення даних в реальному часі, отримавши, таким чином, першу систему доповненої реальності, що носить. Наприкінці 1980-х років інженер із Центру досліджень НАСА Скот Фішер скомбінував відеошолом та рукавички управління віртуальною реальністю, створивши архетипічний інтерфейс [11].

Виникнення та розвиток доповненої реальності. Сам термін «доповнена реальність» виник 1990 року. Цей термін запропонували два інженери дослідницької лабораторії корпорації Boeing: Том Коделл та Девід Майзелл, які працювали над проблемою оптимізації робочого процесу збиравців літаків. Робітникам доводилося постійно звірятися з кресленнями, що уповільнювало темп виконання. Так і народилася думка оснастити співробітників компанії спеціальними пристроями, які б накладали на видимі очима області простору віртуальні об'єкти, написи та схеми, що надають необхідну інформацію в реальному часі. Таким чином, робітники могли бачити, який кабель де треба прокласти, які деталі і як з'єднати. Це призвело б до економії часу, який потрібний для звернення до паперових схем, але розвиток технологій на той момент не дозволив досягти прийнятної компактності та зручності використання, тому від ідеї відмовилися.

Термін «доповнена реальність» позначає один із видів змішаної реальності, в якій зображення реальних об'єктів доповнено віртуальними елементами. Обов'язковою умовою доповненої реальності є те, що

«з'єднання» реального та віртуального світів відбувається у режимі реального часу. У 1994 році професор університету Торонто Пол Мілгрем та професор університету Осаки Фуміо Кішино описали континуум реальність-віртуальність (RV-continuum). Це простір між реальністю та віртуальністю, де розташовані доповнена реальність (ближче до реальності) та доповнена віртуальність, що ближче до віртуальності.

Також Мілгрем і Кішино ввели термін "змішана реальність", що включає в себе "...все між крайнощами віртуального континууму", де віртуальний континуум поширений від повної реальності до повністю віртуального оточення з доповненими реальністю та віртуальністю всередині [12]. Цей континуум складається з двох осей Медіальної реальності концепту Манна, яка реалізується.

Цей континуум складається з двох осей Медіальної реальності концепту Манна, яка реалізується шоломами і комп'ютерами, що носяться, і фотографічними системами, створеними ним у 1970-х і на початку 1980-х, друга вісь є медіальним континуумом, який включає, наприклад, знижену реальність, яка імплементована у шоломах чи окулярах, які блокують рекламу чи замінюють її корисну інформацію (рисунок 1.1) [1, 3, 5].

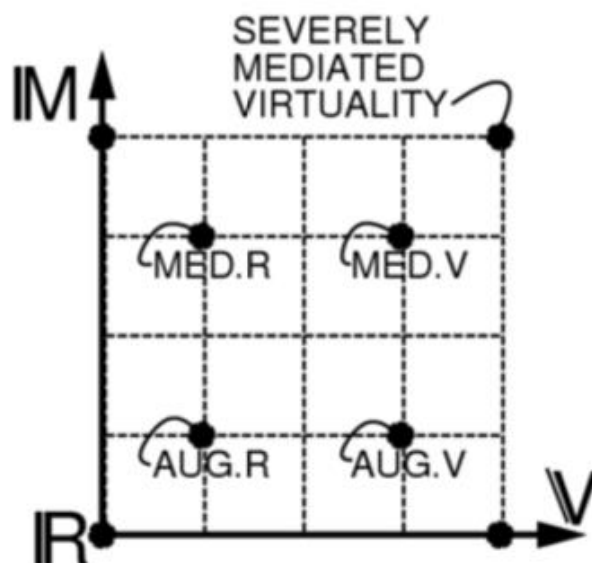


Рисунок 1.1 – Континуум медіальної реальності

На рисунку 1.1 зображено вісь віртуальності (ліворуч-праворуч) і вісь медіальності (знизу-вгору) континууму опосередкованої реальності. Тут показано чотири приклади точок: доповнена реальність, доповнена віртуальність, опосередкована реальність та опосередкована віртуальність на віртуальній та медіальній осях. Це включає, наприклад, знижену реальність (тобто комп'ютеризований шолом, який відфільтровує або знижує окремі частини сцени).

Доповнена реальність на даний час. Вперше об'єднати віртуальні об'єкти і зображення з камери вдалося Хіроказ Като в 1999 році, який створив спеціальну бібліотеку програмного забезпечення з відкритим вихідним кодом ArToolKit [16]. Саме з цього моменту технологія доповненої реальності набула масового поширення. У 2005 році була випущена версія бібліотеки для смартфонів під управлінням операційної системи Symbian (вона не набула широкого поширення, тому що програми, створені з її допомогою, демонстрували низьку продуктивність), в 2008 році для iOS і, нарешті, для Android в 2010 У 2008 році рівень розвитку технологій досяг такого рівня, що обчислювальні потужності смартфонів змогли забезпечити взаємодію з реальними об'єктами як реального часу. Після цього стали виникати загальнодоступні мобільні програми з використанням технології доповненої реальності і як наслідок - повсюдне поширення технології. Використання смартфонів як апаратної платформи для систем доповненої реальності дозволяє суттєво скоротити вартість розробки подібних систем, тому що не доводиться розробляти вже існуючу апаратну частину. У 2012 році компанія Google представила окуляри Google Glass, які використовують доповнену реальність для виведення інформації поверх зображення реального світу.

### 1.3 Класифікація існуючих відеоінформаційних систем доповненої реальності

Людина отримує уявлення про навколишній простір за допомогою великого набору органів чуття. Системи доповненої реальності, здійснюючи ефект занурення в імерсійне середовище, створюють сигнал одного або кількох таких органів. Таким чином, за типом переданої людині інформації системи доповненої реальності мають наступні класифікації: візуальні, аудіо, аудіовізуальні.

**Візуальні.** У основі лежить зорове сприйняття людини. Завдання таких систем – створити зображення, яке буде використано людиною. Оскільки зображення для людини є більш інформативним та зрозумілим, такий вид систем є найпоширенішим.

**Аудіо.** Такі системи орієнтовані на слухове сприйняття та використовуються у навігації [1, 7]. Наприклад, вони видають спеціальні сигнали, коли людина досягає певного місця. Можливе використання стереоскопічного ефекту, що дозволяє людині йти у потрібному напрямку, орієнтуючись на джерело звуку.

**Аудіовізуальні.** Це комбінація двох попередніх типів, проте аудіоінформація у них має лише допоміжний характер. Системи доповненої дійсності потребують інформації, одержуваної з довкілля, т.к. тільки на основі цих даних можна забезпечити взаємодія із реальними об'єктами.

Кожна з таких систем має певний набір пристроїв-приймачів, що дозволяють збирати інформацію з навколишнього середовища: відеоряд, звукові та електромагнітні коливання, прискорення, положення та орієнтація пристрою у просторі тощо.

Розглянемо класифікації за типом приймачів системи доповненої дійсності.

**Геопозиційні.** Орієнтуються насамперед на сигнали систем глобального географічного позиціонування GPS або ГЛОНАСС. Крім приймачів таких

сигналів геопозиційні системи можуть використовувати магнітний компас, гіроскоп і акселерометр для більш точного визначення положення та орієнтації пристрою в просторі.

Оптичні (відеоінформаційні) системи обробляють відеоряд, отриманий з камери, і на підставі аналізу відеоряду доповнюють зображення реального світу віртуальними об'єктами.

Комбіновані системи використовують одночасно як оптичні дані, так і дані геопозиціонування, що дозволяє досягти високої точності визначення положення системи доповненої реальності у просторі.

Оптичні системи доповненої реальності можна поділити на способи отримання інформації про реальний світ. Відповідно до такого поділу, системи можна поділити на маркерні та немаркерні.

Маркерні системи історично з'явилися першими, т.к. є найпростішими у реалізації, але водночас найпримітивнішими. Маркерні системи доповненої реальності здійснюють пошук заздалегідь заданих спеціальних маркерних зображень (рисунок 1.2) у відеопотоці, що входить. Потім на основі інформації про положення та орієнтацію маркерів система виконує ту чи іншу дію [1,8].



Рисунок 1.2 – Маркерне зображення

Безмаркерні системи реалізують більш складні алгоритми обробки зображень, і для визначення положення та орієнтації в просторі використовують не спеціальні маркерні, а звичайні зображення, які мають певні достатні характеристики (рисунок 1.3). Найважливішою характеристикою таких зображень є кількість унікальних контрастних областей [1,9].

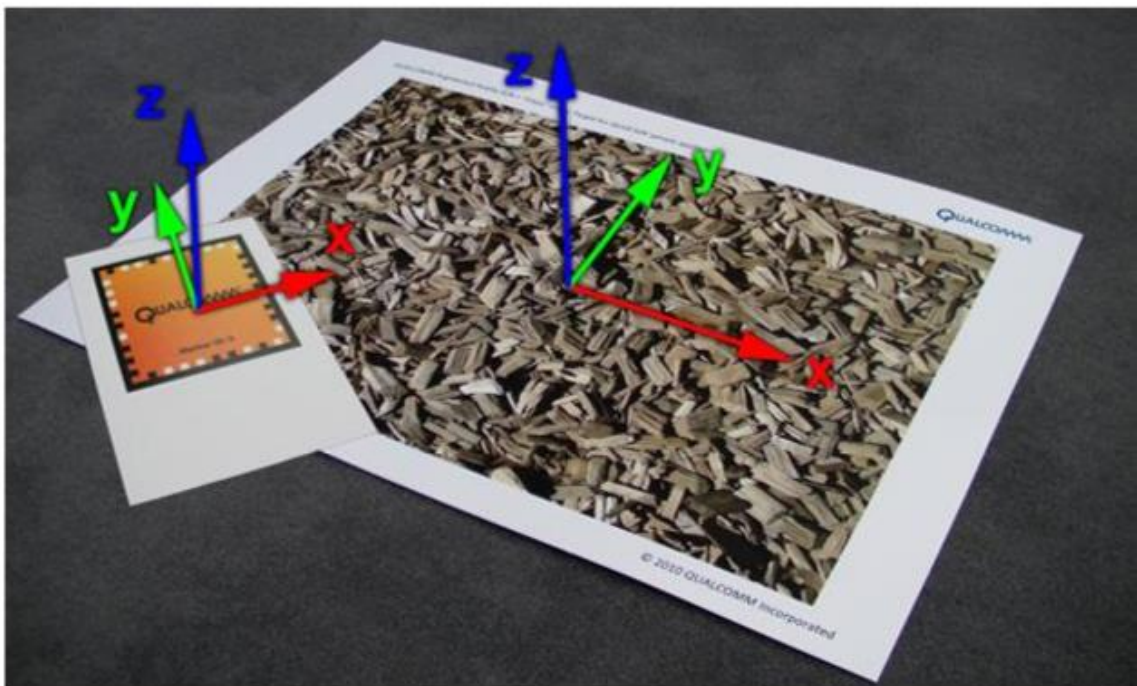


Рисунок 1.3 – Приклади зображень, що використовуються у безмаркерних системах доповненої реальності

Системи просторового трекінгу не потребують спеціальних зображень. Поточне положення у просторі визначається в режимі реального часу на основі безперервного аналізу відеопотоку, що надходить із відеокамери [2]. Системи доповненої реальності можна розрізняти за рівнем взаємодії з користувачем.

У деяких системах користувач грає пасивну роль, він лише спостерігає за реакцією системи на зміни у навколишньому середовищі. Інші системи потребують активного втручання користувача – він може керувати як роботою самої системи для досягнення результатів, так і змінювати

віртуальні об'єкти. За цією ознакою системи діляться в такий спосіб.

Автономні. Вони не потребують втручання користувача. Завдання таких систем зводиться до надання інформації про об'єкти. Наприклад, такі системи можуть аналізувати об'єкти, що у полі зору людини і видавати довідкову інформацію про них. Також системи такого типу використовуються у медицині.

Інтерактивні. Такі системи ґрунтуються на взаємодії з користувачем, який отримує відповідь системи на різні дії. У подібних системах необхідний пристрій введення інформації, і як такий пристрій може використовуватися сенсорний екран мобільного телефону, планшета або спеціальний маніпулятор.

Або, як і з нашлемними системами індикації, пристроєм введення може бути підсистема, яка відстежує рух очей. Вибір пристроїв введення залежить від специфіки системи. У разі простих дій з віртуальним об'єктом досить простого вказівного пристрою. Якщо ж потрібна імітація будь-яких реальних процесів та виконання складних маніпуляцій з об'єктами, використовуються спеціальні маніпулятори, що мають різну кількість ступенів свободи. Інтерактивність систем доповненої реальності виявляється різною мірою. Існують системи, що дозволяють користувачеві активно змінювати віртуальне середовище.

Зазвичай це системи-симулятори будь-яких реальних дій. Вони застосовуються у разі, коли використання реальних об'єктів неможливо, наприклад, спеціалізовані медичні тренажери, що дозволяють початківцям відпрацьовувати необхідні навички. Існують інші системи, де користувачеві не потрібно змінювати віртуальне середовище. Натомість користувач вибирає, які віртуальні об'єкти він хоче побачити. Користувач також має можливість маніпулювати віртуальними об'єктами, але не лише на рівні структури, але в рівні відображення, тобто. застосовувати, наприклад, афінні перетворення: поворот, рух тощо. До цієї групи можна віднести різні архітектурні системи, що дозволяють побачити, як впишеться в реально

існуючу обстановку нову споруду або її частину, а також навігаційні та геоінформаційні системи. Подібні системи можуть показувати частини об'єктів інтересу, приховані іншими спорудами, додаткову інформацію про вибрані об'єкти тощо.

За рівнем мобільності системи доповненої реальності можна класифікувати на стаціонарні та мобільні.

Стаціонарні системи призначені до роботи у фіксованому місці; переміщення таких систем означає часткове або повне припинення їхньої працездатності.

Мобільні системи можуть легко переміщатися; часто таке переміщення і є основою виконуваної ними функції. До них належать відеоінформаційні системи доповненої дійсності, які працюють на смартфонах чи планшетних комп'ютерах, тобто. переважна більшість існуючих систем. На рисунку 1.4 наведено фотографію прототипу однієї з перших мобільних систем доповненої реальності.



Рисунок 1.4 – Прототип однієї з перших мобільних систем доповненої реальності

У системах доповненої реальності може використовуватись автоматизована інформаційна система, що містить інформацію про об'єкти реального світу. Складові такої системи класифікуються на стаціонарну частину та мобільну частину.

Стаціонарна частина може представляти віддалений сервер з базою даних про реальні об'єкти та їх місцезнаходження.

Мобільна частина являє собою пристрій, що носить, який може включати в себе модуль геолокації.

Мобільна частина отримує дані про своє місцезнаходження та орієнтацію та відсилає ці дані на віддалений сервер, де вони використовуються як умова вибірки з бази даних. Потім віддалений сервер надсилає на мобільну частину інформацію про реальні об'єкти, які потрапляють у поле зору користувача.

Такий підхід може бути застосований у навігаторах чи геоінформаційних системах [2]. Приналежність до того чи іншого типу визначається функціями системи. Так, симулятор хірургічного столу не повинен бути мобільним, оскільки його завдання – відтворити для людини спеціальні умови, які максимально наближені до реальних. У той же час навігаційна система має бути якомога мобільнішою, щоб вона могла переміщатися разом з транспортним засобом або людиною, не створюючи додаткових витрат на її переміщення.

## 2 РОЗРОБКА МЕТОДУ ПІДВИЩЕННЯ ШВИДКОДІЇ МОБІЛЬНИХ ВІДЕОІНФОРМАЦІЙНИХ СИСТЕМ ДОПОВНЕНОЇ РЕАЛЬНОСТІ

### 2.1 Особливості відеоінформаційних систем доповненої реальності

Основне завдання під час реалізації системи систем доповненої реальності – визначити взаємне розташування особливого «маркерного» зображення та спостерігаючої камери або зміна її положення у просторі за допомогою аналізу відеоряду [2].

В даний час багато мобільні пристрої мають вбудоване апаратне забезпечення для визначення положення пристрої у просторі. До такого апаратного забезпечення відносяться гіроскоп, акселерометр, магнітний компас. Використовуючи інформацію, отриману від зазначеного апаратного забезпечення, можна визначати становище та орієнтацію мобільного пристрої у просторі у будь-який момент часу [2, 3]. Таким чином, точність визначення зміни положення геопозиційної системи доповненою реальності у просторі обмежується точністю геоприймачів, гіроскопів та іншого апаратного забезпечення для визначення положення у просторі. З цього слід висновок, що однієї із завдань технології доповненою реальності є підвищення точності геопозиціонування та визначення орієнтації у просторі.

Для оптичних систем доповненої реальності характерні інші складності, що полягають в алгоритмах обробки зображень, а також у підвищенні якості оброблюваних зображень. Кадр відеоряду, отриманий з камери, що використовується в системі, може не мати необхідною чіткістю, умови навколишнього середовища також можуть негативно позначитися на якості отриманого зображення. Завдання оптичних систем доповненої реальності можуть бути поділені на дві категорії: завдання покращення апаратної складника та завдання покращення програмного забезпечення. У загальному вигляді алгоритм визначення зміни положення спостерігаючою відеокамери наведено у блок-схемі на рисунку 2.1.

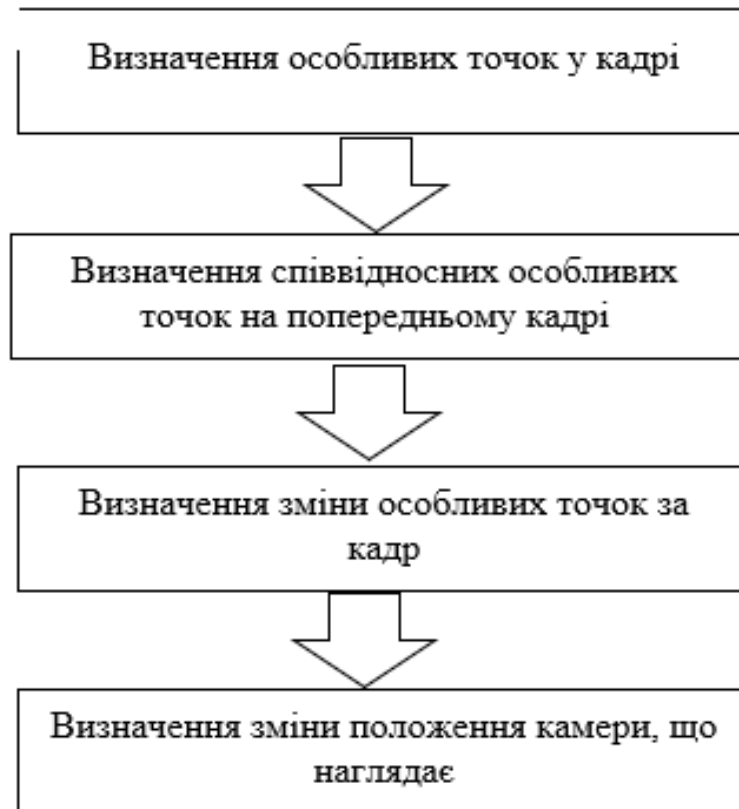


Рисунок 2.1 – Блок-схема алгоритму визначення зміни положення відеокамери

Для вирішення даної завдання у кожному кадрі відеоряду, що надходить з камери, визначаються особливі точки зображення (локальні особливості). Особлива точка зображення – це точка, околиця якої можна, можливо відрізнити від околиці будь-якої іншої точки зображення [2, 4].

## 2.2 Визначення локальних особливостей зображення

Локальна особливість зображення – це така точка зображення, околиця якої можна, можливо відрізнити від околиці будь-якої іншої точки зображення. Відділення хибних особливостей від справжніх є однією з основних проблем під час відстеження русі камери або при реконструкції тривимірний сцени. Погіршується вона тим, що надійно вирішити її, працюючи тільки з одним зображенням неможливо. Необхідно встановлювати відповідність між особливостями кількох зображень, а потім

побудувати деяку модель, якою ці відповідності будуть задовольняти найкращим чином. Особливі точки, відповідності яких не будуть задовольняти такий моделі, будуть визнані помилковими, або викидами. Методи побудови таких моделей в останнє десятиліття бурхливо розвивалися. Алгоритми, про які розповідається далі, застосовні для всіх точкових особливостей зображень. Проте необхідно додатково відзначити, що при стеженні за хибною точковою особливістю зображення на різних кадрах послідовності їй будуть відповідати різні точки сцени. Для простоти як околиці точки зображення береться прямокутне вікно невеликого розміру. Для порівняння таких прямокутних вікон можуть використовуватися різні заходи на зображеннях (наприклад, кроскореляція).

Більшість детекторів точкових особливостей працюють подібним чином: для кожної точки зображення обчислюється деяка функція від неї околиці. Крапки, в яких ця функція досягає локального максимуму, очевидно, можна відрізнити від усіх точок з деякою її околиці. Існує великий набір функцій, які можна, можливо використовувати для виявлення точкових особливостей. Найчастіше всього для завдань відстеження точок сцени застосовуються функції, що у зображенні структури, схожі на кут. Детектори, що використовують такі функції, що називаються детекторами кутів. Саме вони частіше всього застосовуються для вирішення завдань відстеження точкових особливостей сцени. Кути – особливі точки, які формуються з двох або більше граней, і грані зазвичай визначають кордон між різними об'єктами або частинами того самого об'єкта [5]. Інакше говорячи, кут – це точка, у якій в околиці інтенсивність змінюється щодо центру. Кути визначаються за координатами та змінами яскравості навколишніх точок зображення. Головна властивість кутів полягає в тому, що в області навколо кута у градієнта зображення переважають два домінуючі напрями, що робить їх помітними. Градієнт – векторна величина, що показує напрямок найшвидшого зростання (зменшення) функції інтенсивності зображення. Так як зображення дискретно, то вектор градієнта визначається через приватні

похідні по осі  $X$  та  $Y$  через зміни інтенсивностей сусідніх точок зображення . Більшість методів розглядають незграбність, що залежить від похідної другого порядку , тому загалом методи чутливі до шуму.

Існують різні види кутів : L-, Y- ( або T-), і X- зв'язкові (деякі виділяють ще стрілоподібно зв'язкові кути ) [4, 5]. приклади різних видів кутів наведено на рисунку 2.2 .



Рисунок 2.2 –Види кутів

Існує три основні підходу до визначення локальних особливостей зображення [6].

Засновані на інтенсивності зображення : особливі точки обчислюються безпосередньо із значень інтенсивності пікселів зображення .

Використовують контури зображення: методи вилучають контури та шукають місця з максимальним значенням кривизни або роблять апроксимацію контурів та визначають перетину. Такі методи чутливі до околиць перетинів , оскільки вилучення часто може бути неправильним у тих місцях , де перетинаються 3 або більше країв.

На основі використання моделі: використовуються моделі з інтенсивністю як параметрів, які підлаштовуються до зображень -шаблонів до субпіксельної точності, мають обмежене застосування з особливими точками спеціальних видів (наприклад, L- зв'язковими кутами) і залежать від використовуваних шаблонів.

На практиці для широкого застосування найбільш поширені методи, засновані на інтенсивності зображення [7].

### 2.3 Детектування особливих точок

Детектор Моравця. Робота у дослідженні прив'язки зображень з використанням особливих точок почалася детектора Моравця [8]. Детектор Моравця – самий простий з існуючих. Автор розглядає зміна яскравості квадратного вікна (звичайно розміру 3x3, 5x5, 7x7 пікселів) цікавої точки при зрушенні вікна  $W$  на 1 піксель у восьми напрямках (горизонтальних, вертикальних і діагональних) [2, 9]. Основними недоліками розглянутого детектора є відсутність інваріантності до перетворення повороту та виникнення помилок детектування за наявності великого кількості діагональних ребер [3].

Детектор Харріса. Як показали дослідження, найбільш оптимальним детектором  $L$  зв'язкових кутів є широко відомий детектор Харріса (також його називають оператором Плесея або детектором Харріса і Стефенса) [3]. Харріс і Стефенс покращили детектор Моравця, ввівши анізотропію по всіх напрямів, тобто. використовували похідні яскравості зображення для дослідження змін яскравості за безліччю напрямків.

Для зображення  $I(x, y)$  розглянемо прямокутну область  $W(x, y)$ , що є околиці точки  $C(x, y)$ , де  $(x, y)$  – координати центру околиці,  $u$  – зсув строка по горизонталі,  $v$  – зсув строка по вертикалі. Зазвичай розмір області  $W$  дорівнює 5x5 пікселів, але він може залежати і від розміру зображення, отже в більшості випадків  $(u, v) \in [-2; 2]$ . Зважена сума квадрата різниць між зрушеної та вихідної областю (тобто зміна околиці точки  $(x, y)$  при зрушенні на  $(u, v)$ ) дорівнює

$$E(x, y) = \sum_{(u, v) \in W} (I(x + u, y + v) - I(x, y))^2 \approx [x \quad y] \mathbf{M} \begin{bmatrix} x \\ y \end{bmatrix},$$

де  $\mathbf{M}$  – автокореляційна матриця:

$$\mathbf{M} = \sum_{(u,v) \in W} \begin{pmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{pmatrix}, \quad (2.1)$$

а  $I_x$  та  $I_y$  – приватні похідні яскравості по рядку та стовпцю, відповідно. Кут характеризується великими змінами функції  $E(x, y)$  по всьому можливим напрямкам  $(x, y)$ , що еквівалентно великим за модулем власним значенням матриці  $\mathbf{M}$ .

Оскільки безпосередній розрахунок власних значень є трудомісткою завданням, Харрісом і Стефеном була запропоновано міру відгуку

$$r = |\mathbf{M}| - k[\text{tr}(\mathbf{M})]^2, \quad (2.2)$$

де  $k$  – емпірична константа,  $\text{tr}(\mathbf{M})$  – слід матриці. Крапки, у яких значення  $r$  більше порогового, вважаються особливими.

Детектор Харріса інваріантний до поворотів, частково інваріантний до афінних змін інтенсивності. До недоліків стоїть віднести чутливість до шуму та залежність детектора від масштабу зображення [3].

Детектор Shi-Tomasi. Цей детектор багато в чому збігається з детектором Харріса, але відрізняється у обчисленні заходи відгуку: алгоритм безпосередньо обчислює мінімальне значення з власних чисел автокореляційної матриці  $\mathbf{M}$  (2.1)  $\min$ , замість заходи відгуку (2.2) оскільки робиться припущення, що пошук кутів буде більше стабільним [3].

Детектор SUSAN. Алгоритм SUSAN (Smallest Univalued Segment Assimilation Nucleus) був запропонований Смітом та Бреді [3,4]. Кути в цьому алгоритмі визначаються сегментацією кругових околиць у схожі (помаранчеві) і несхожі (сині) ділянки (рисунки 2.3). Кути знаходяться там, де відносна площа подібних ділянок досягає локального мінімуму нижче певного порога.

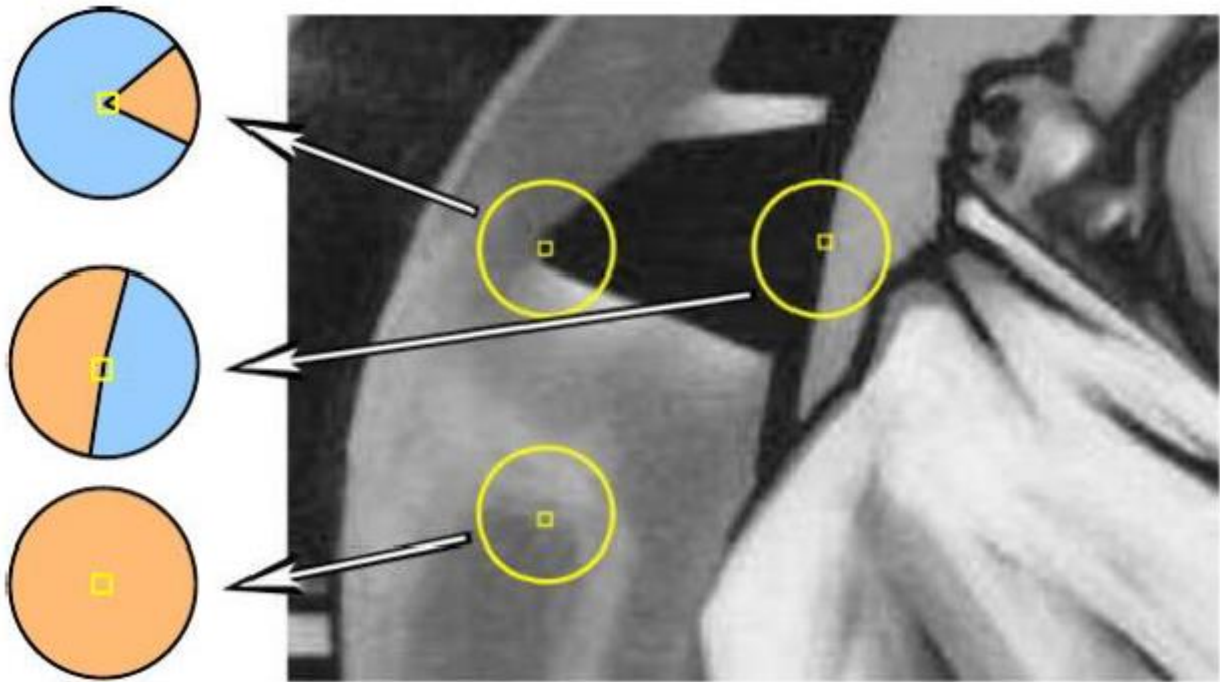


Рисунок 2.3 – Кругові околиці

Алгоритм показує гарну точність до всіх видів кутів, але нестійкий до розмиття на зображеннях [4].

Детектор FAST. Ростен і Драммонд опублікували опис FAST Features from Accelerated Segment Test). В алгоритмі розглядається коло з 16 пікселів, яка виділяється алгоритмом Брезенхема довкола точки-кандидата  $P$  (рисунок 1.10). Крапка є кутовий, якщо для поточної розглянутої точки  $P$  існує  $N$  суміжних пікселів на колі, інтенсивності яких більше  $I_{P+t}$  або інтенсивності всіх менше  $I_{P-t}$ , де  $I_P$  – інтенсивність точки  $P$ ,  $t$  – порогова величина. Далі необхідно порівняти інтенсивність у вертикальних та горизонтальних точках на колі під номерами 1, 5, 9 та 13 з інтенсивністю в точці  $P$  (для того, щоб як можна, можливо швидше відсікти хибні кандидати). Якщо для трьох з цих точок виконається умова  $I_{P_i} > I_{P+t}$  або  $I_{P_i} < I_{P-t}$ ,  $i = 1, \dots, 4$ , то проводиться повний тест для всіх 16 точок. Експерименти показали, що найменше значення  $N$ , при якому особливі точки починають стабільно виявлятися  $N = 9$  [5].

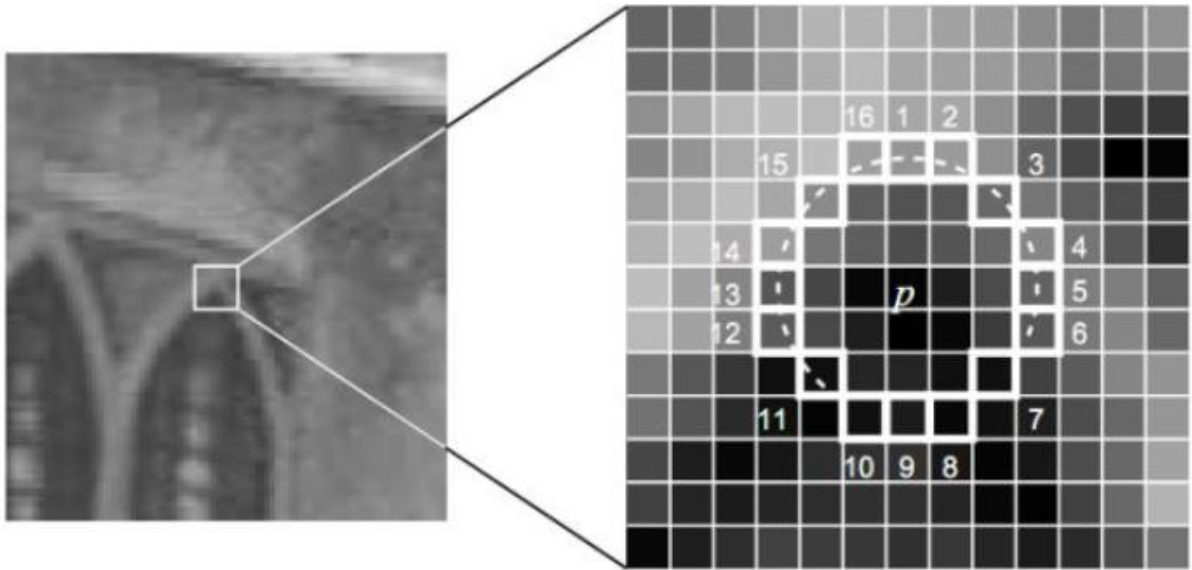


Рисунок 2.4 – Виділення кола в алгоритмі FAST

Алгоритм має ряд недоліків, наприклад, поблизу деякою околиці може виявитися кілька особливих точок, та ефективність алгоритму залежить від порядку обробки зображення та розподілу пікселів.

Порівняння поширених детекторів особливих точок. Нижче наведено порівняльну таблицю (таблиця 2.1) для описаних детекторів кутів. В якості критеріїв порівняння були обрані ефективність виявлення, повторюваність, стійкість до спотворень та швидкість. Порівняння критеріїв проводилося за п'ятибальною шкалі (1 – дуже погано, 2 – погано, 3 – задовільно, 4 – добре, 5 – відмінно).

Як слід з таблиці, найбільш ефективні детектор Харріса та його варіація (детектор Shi-Tomasi) та SUSAN-детектор. З точки зору повторюваності SUSAN-детектор показує незадовільні результати через нестійкість до афінних перетворень, які повсюдно зустрічаються при розгляді кадрів відеоряду. З точки зору швидкодії детектор Shi-Tomasi програє детектору Харріса при подібній ефективності.

Таблиця 2.1 – Порівняння детекторів кутів

Алгоритм	Ефективність виявлення	Повторюваність	Стійкість до спотворень	Швидкість
Детектор Моравця	3	3	3	4
Детектор Харріса	4	5 – афінні перетворення, 3 – масштабування	3	2
Детектор Shi-Tomasi	4	5 – афінні перетворення, 3 – масштабування	3	1
Детектор SUSAN	4	4 - масштабування, 2 – афінні перетворення	5	4
Детектор FAST	3	4	3	4

На основі порівняння вищеписаних детекторів та з урахуванням описаних недоліків детектор Харріса бачиться як найбільш перспективний під час реалізації відеоінформаційних систем доповненої реальності, але для ефективного використання вказаного детектора потрібно підвищити його швидкодію.

#### 2.4 Вимоги до дескрипторів локальних особливостей зображення

Кореляційний метод, який застосовується у фотограмметрії для порівняння фрагментів зображень при малих геометричних спотвореннях, дуже чутливий до афінних спотворень, повороту та масштабування. Тому для опису локальних околиць характерних точок конструюються спеціальні дескриптори, що мало змінюються при різних перетвореннях зображень.

Сформульовано наступний ряд вимог до дескрипторів локальних особливостей зображення [3, 6].

1 Інваріантність – дескриптори однієї і тієї ж точки або області, що лежить на двох різних зображеннях, не повинні сильно відрізнитися.

2 Унікальність – дескриптори двох різних особливостей зображення

повинні помітно відрізнятись один від одного.

3 Стійкість – дескриптор однієї і тієї ж точки не повинен сильно змінюватися при геометричних перетвореннях зображення.

У реальній ситуації дескриптори не мають усіма зазначеними властивостями одночасно, і доводиться вибирати той дескриптор, який найкраще відповідає Розв'язуване завдання. У найпростішому випадку дескриптори обчислюються за круглою околиці особливої точки. Її радіус або фіксується заздалегідь, або для кожної особливої точки обчислюється динамічно, наприклад, як величина, що доставляє максимум деякої функції інформативності, порахованої за фрагментом зображення [7, 8].

## 2.5 Побудова швидкісних дескрипторів особливих точок

На початку 2000-х років почали розвиватися непараметричні методи описи локальних особливостей, засновані на гістограмах різних властивостей зображення, таких як яскравість та її похідні, кривизна, відгуки зображення на вплив наборами фільтрів різної частоти та спрямованості.

Дескриптор SIFT. Найбільш найпопулярнішим дескриптором на даний момент є дескриптор SIFT [3, 9]. Цей дескриптор представляє собою гістограму напрямків градієнтів зображення на околиці особливої точки. Околиця локальною особливості зображення ділиться на чотири квадратний сектор. В кожному пікселі всередині кожного сектора обчислюється градієнт зображення, його напрямок та модуль (рисунок 1.11). Потім модулі градієнтів множаться на вагу, експоненційно спадаючий з віддаленням від точки інтересу. Це зроблено для того, щоб уникнути різких змін значення дескриптора при невеликих змінах положення вікна, а також для того, щоб градієнти, віддалені від центру дескриптора, вносили менший внесок у його значення, оскільки градієнти на периферії околиці точки інтересу найменш стійкі при геометричних перетвореннях зображення. По кожному сектору збирається гістограма напрямів градієнтів, причому кожне входження зважується модулем градієнта.

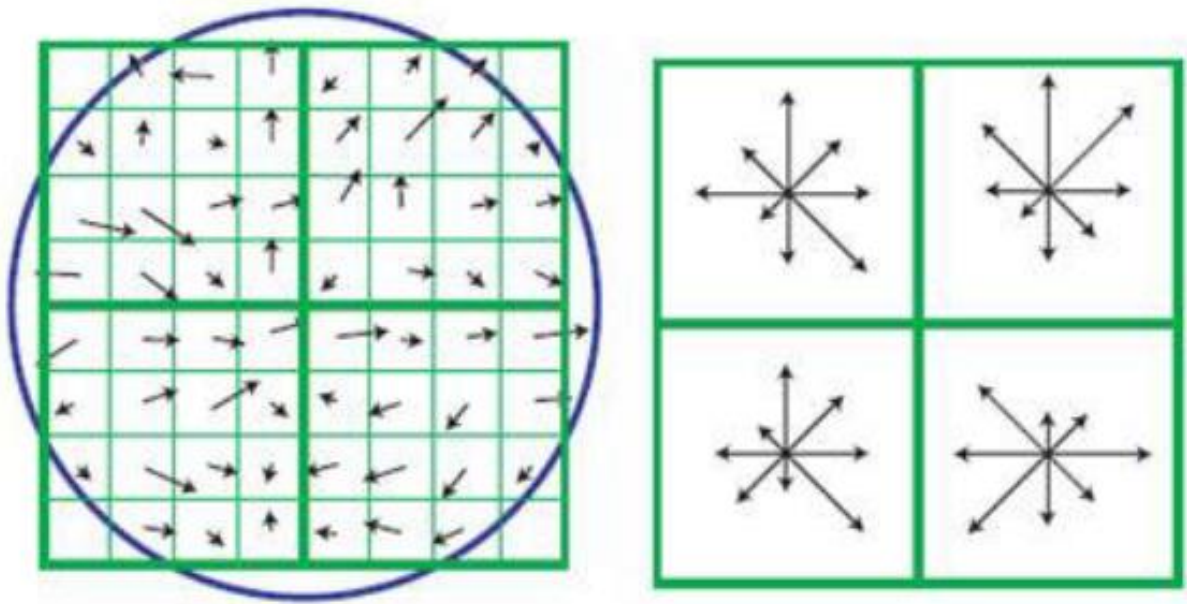


Рисунок 2.5– – Побудова дескриптора SIFT

Дескриптор SIFT представляє собою вектор, отриманий зі значень усіх елементів гістограм напрямків, і складається із 128 компонент. Потім дескриптор нормується, щоб підвищити його стійкість до змін яскравості. Дескриптор SIFT широко застосовується у комп'ютерному зорі. На ідеї використання локальних гістограм градієнтів яскравості зображення засновано більшість сучасних дескрипторів. Відмінності більше нових дескрипторів полягають у використанні осередків у формі кругових секторів та кілець (наприклад, дескриптори GLOH та DAISY) [7].

Дескриптор SURF. У 2008 був представлений найближчий конкурент SIFT-дескриптора, SURF-дескриптор [4]. Процедура побудови SURF-дескриптора схожа на SIFT, але в ній використовуються не гістограми зважених градієнтів, а відгуки вихідного зображення на вейвлети Хаара. На першому кроці отримання дескриптора навколо особливої точки будується квадратна область, яку орієнтують по деякому переважному напрямку. Потім область поділяється на квадратні сектори. В кожному з секторів у точках, що належать регулярною сітці, обчислюються відгуки на два види вейвлетів

Хаара – горизонтально та вертикально спрямовані (рисунок 2.6). Відгуки зважуються гауссіаном, що підсумовуються по кожному сектору, і утворюють першу частину дескриптора.



Рисунок 2.6 – Вейвлети Хаара та області обчислення

Друга частина процедури побудови дескриптора складається з побудови гістограми сум модулів відгуків, що вказують на структуру зображення. На рисунку 2.7 зліва направо: однорідна область, анізотропна текстура, градієнтна текстура. SURF-дескриптор має довжину 64 числа. Як і SIFT, SURF-дескриптор інваріантний до адитивного зміни яскравості. Інваріантність до мультиплікативного зміни яскравості досягається шляхом нормування дескриптора.

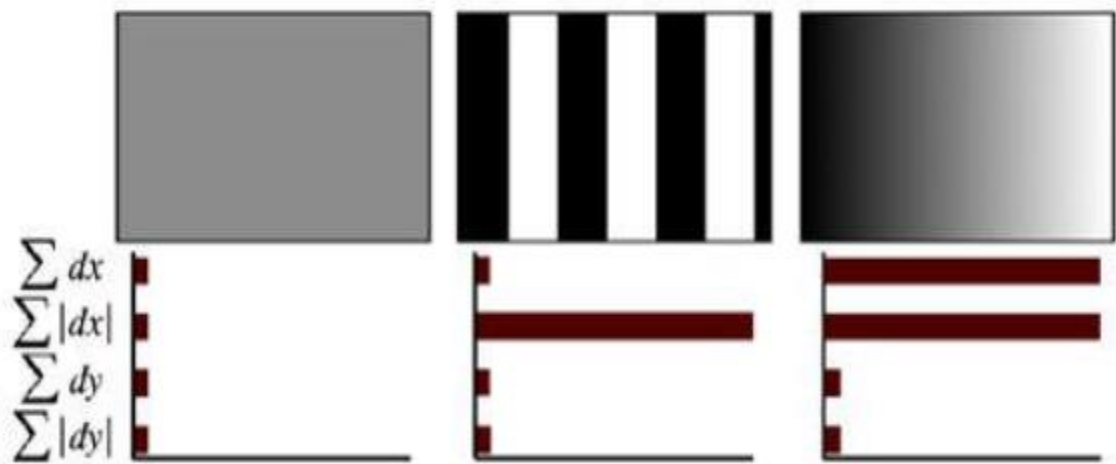


Рисунок 2.7 – Гістограми сум модулів відгуків

Серед алгоритмів виділення особливих точок існує поділ на більш математичні обґрунтовані, але відносно повільні (детектор Харріса, детектор Shi-Tomasi) і більше евристичні, але швидкі (SUSAN, FAST). Той самий поділ застосовно і до дескрипторів локальних особливостей зображення тільки тут як роздільної ознаки виступає компактність та простота обчислення дескриптора. Чим менше довжина дескриптора, тим менше пам'яті потрібно для його зберігання, і менше часу порівняння його з іншими.

## 3 ДОСЛІДЖЕННЯ МЕТОДУ ПІДВИЩЕННЯ ШВИДКОДІЇ МОБІЛЬНИХ ВІДЕОІНФОРМАЦІЙНИХ СИСТЕМ ДОПОВНЕНОЇ РЕАЛЬНОСТІ

### 3.1 Оптимізація детекторів локальних особливостей зображення

Відеоінформаційні системи доповненою реальності (ВІС ДР) класифікуються за способом отримання інформації про реальний світ. Виділяють маркерні, безмаркерні системи, а також системи на основі просторового трекінгу.

Безмаркерні системи доповненою реальності, а також системи на основі просторового трекінгу використовують два основні підходу до визначення положення спостерігаючої камери в просторі, метод виділення локальних особливостей зображення та метод порівняння текстур.

Метод виділення локальних особливостей зображення використовує детектори локальних особливостей, які в основному побудовані на обчислювально-ємних алгоритмах. Особливість мобільних відеоінформаційних систем доповненої реальності полягає в тому, що, з одного боку, такі системи мають обмеженими обчислювальними ресурсами, що викликано прагненням мінімізувати енергоспоживання мобільного пристрою в цілому, а з іншого боку, відеоінформаційні системи доповненою реальності повинні працювати у режимі реального часу [3]. Ці особливості накладають обмеження на використовувані алгоритми і посилюють вимоги щодо швидкодії, що пред'являються до них.

Цей розділ присвячений оптимізації детекторів локальних особливостей зображення (ДЛОЗ) для їх наступного застосування в мобільних відеоінформаційних системах доповненої реальності.

У відеоінформаційних системах доповненою реальності, які використовують принцип просторового трекінгу, основне завдання полягає в тому, щоб безперервно визначати зміну положення спостерігаючої камери в просторі [2]. Для вирішення цієї задачі використовується алгоритм,

зображений на рисунку 3.1 .

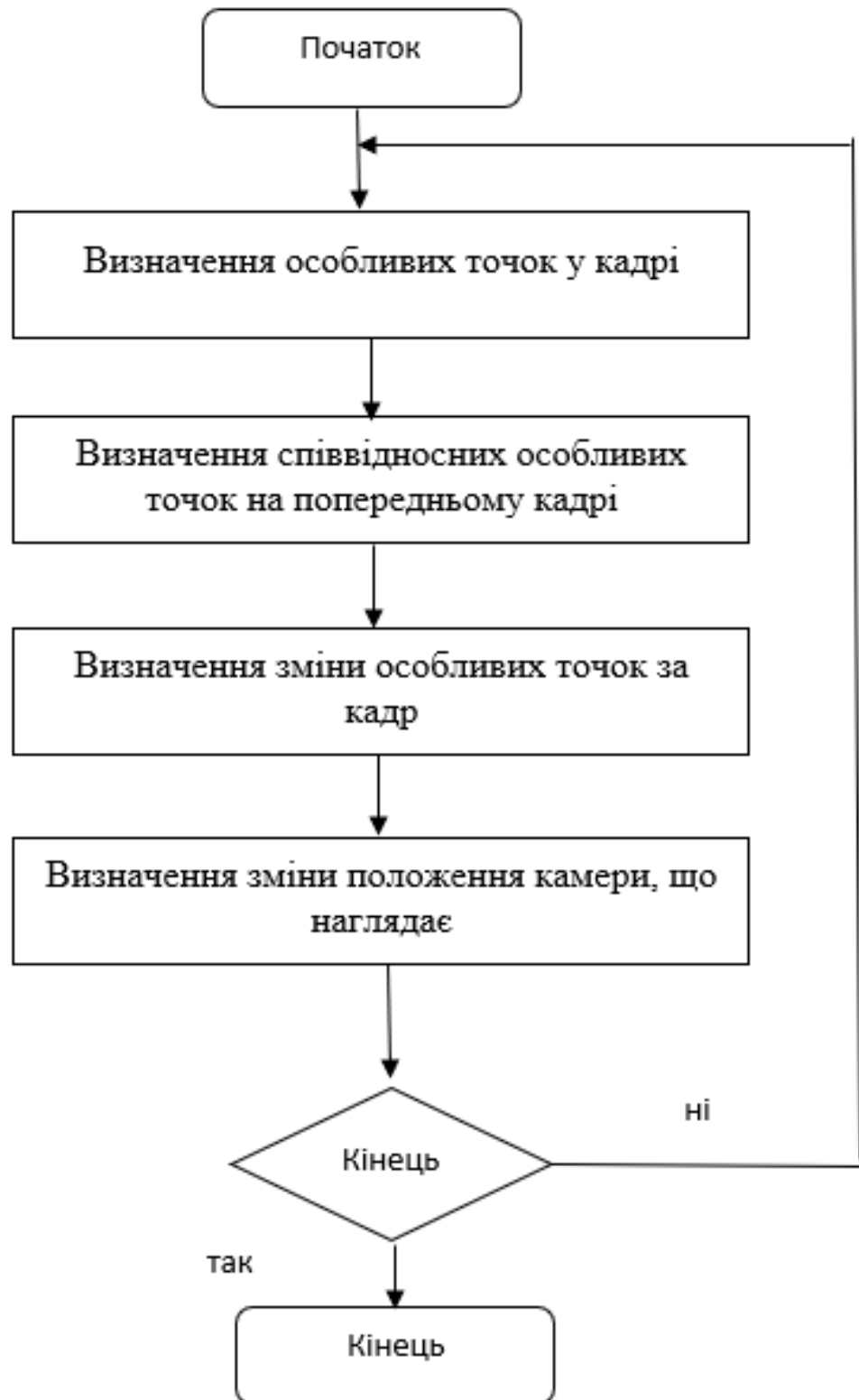


Рисунок 3.1 – Алгоритм методу

В кожному новому кадрі визначаються локальні особливості зображення, які зіставляються з особливостями, знайденими у попередньому кадрі. Далі визначається зміна положення зіставлених особливих точок та на основі цієї інформації розраховується зміна положення спостерігаючої камери в просторі.

Для виділення особливих точок з відеопотоку, що надходить з камери, використовуються спеціальні алгоритми – детектори локальних особливостей зображення, які зазвичай класифікують за типом виділених особливих точок. Найбільш поширені два типи особливих точок: кути та кола (блоби).

Кут – це особлива точка, яка формується з двох і більше граней, грані зазвичай визначають кордон між різними об'єктами або частинами того самого об'єкта [4]. Основна властивість кутів полягає в тому, що в околиці кута у градієнта зображення переважають два напрями, що робить їх помітними. Градієнт – векторна величина, що показує напрямок найшвидшого зростання функції інтенсивності зображення  $I(x, y)$ . Залежно від кількості перетинаються граней існують різні види кутів, зображені на рисунку 3.2: L-, Y- (чи T-), і X- зв'язкові (деякі виділяють ще стрілоподібно зв'язкові кути) [3].



Рисунок 3.2 – Види кутів

У мобільних відеоінформаційних системах доповненої реальності найбільше поширення набули детектори кутів, які мають більше високою ефективністю на відміну блоб -детекторів. Поряд з ефективністю важливою

характеристикою детектора є швидкодія. Так як ВІС ДР є системами реального часу, то для прийнятної якості виведеного відеопотоку необхідно, щоб частота кадрів не опускалася нижче порога 25 кадрів в секунду, що визначає верхню кордон тривалості обробки одного кадру відеопотоку, а саме 40 мс [8].

### 3.2 Метрики оцінки ефективності мобільних відеоінформаційних систем доповненої реальності

Для підвищення швидкодії детекторів кутів пропонується використовувати яскравий попередній детектор особливих точок (ЯПД ВІД). Принцип його дії полягає у виділенні потенційних особливих точок з зображення та передачі їх для подальшого аналізу в детектор. Таким чином, для виділення особливостей буде проаналізовано не весь кадр, а лише ті області, де можливо їх знаходження. Блок-схема алгоритму роботи попереднього детектора наведено на рисунку 3.3.

Для визначення перепадів яскравості будується матриця середньоквадратичною суми приватних похідних яскравості у пікселі

$$S_{x,y} = \sqrt{\left(\frac{dI}{dx}\right)^2 + \left(\frac{dI}{dy}\right)^2}, \quad (3.1)$$

де  $S$  – матриця середньоквадратичних сум приватних похідних яскравості зображення. Після побудови матриці визначаються максимальне та мінімальне значення її елементів та на них основі розраховується порогове значення зміни яскравості

$$d = S_{min} + \left(\frac{S_{max} - S_{min}}{k}\right), \quad (3.2)$$

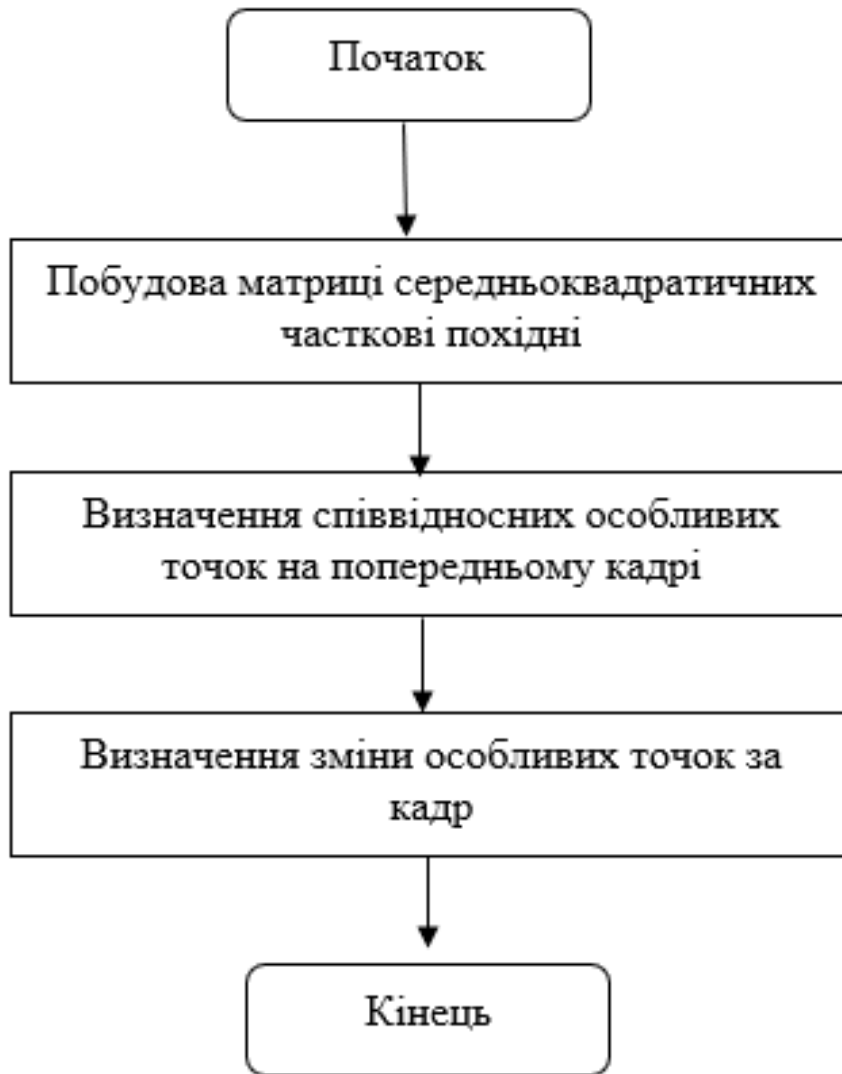


Рисунок 3.3 – Блок-схема основного алгоритму

де  $k$  – емпірична константа, значення якої зазвичай вибирається в діапазоні  $[1, 4 \dots 2]$ ;  $S_{min}$  та  $S_{max}$  – мінімальне та максимальне значення елемента матриці  $S$ , відповідно. потенційної особливою точкою  $I_{x,y}$  вважається така точка зображення  $I$  з координатами  $(x, y)$ , для якої виконується наступне умова :

$$S_{x,y} \geq d. \quad (3.3)$$

У літературі описано достатньо багато різних детекторів кутів : детектори Моравця, Sh-Tomasi, детектор Харріса тощо. Одним з головних недоліків багатьох алгоритмів є їх висока обчислювальна складність при не самій високою точності [4]. Найбільш точні результати детектування

особливих точок показує детектор Харріса , він же є найбільш часто використовуваним детектором локальних особливостей зображення в системах реального часу було проведено дослідження використання попереднього детектора разом із детектором Харріса [4]. Для виділення локальних особливостей зображення детектор Харріса аналізує кожен піксель зображення, а також пікселі на його околиці [9]. Таким чином, для виділення детектором Харріса локальних особливостей зображення , що має розмір  $N$  пікселів у висоту та  $M$  пікселів у ширину (при розмірі околиці  $P \times Q$  пікселів ) знадобиться  $M \times N \times P \times Q$  ітерацій алгоритму. У запропонованому яскравості попередньому детекторі кожен піксель аналізується один раз, тому для попереднього виділення особливих точок потрібно  $M \times N$  ітерацій алгоритму.

Таким чином, використання яскравого попереднього детектора дозволяє з незначними обчислювальними витратами визначити потенційні особливі точки зображення згодом будуть оброблені детекторами кутів , що вимагають значних обчислювальних витрат.

### 3.3 Результати оцінки ефективності запропонованого методу

Для оцінки ефективності застосування попереднього детектора пропонуються наступні метрики. Емуляційний обчислювальний експеримент, що описується далі, вироблявся на комп'ютері з двоядерним процесором AMD Athlon II P360 з тактовою частотою 2,3 ГГц. Зміна тривалості обробки кадру під час використання попереднього детектора. Зміна кількості знайдених особливих точок на зображенні. Зміна порогового значення відгуку кута для набору точок, знайдених детектором Харріса як без використання попереднього детектора особливих точок, так і з його використанням.



Рисунок 3.4– Тестове зображення

Таблиця 3.1 – Тривалість обробки тестового зображення

Роздільна здатність тестового зображення, пікселі	320x240	640x480	768x576	800x600	1024x768
Математичне очікування тривалості обробки без попереднього детектора, мс	1386	5681	7839	8439	14006
Дисперсія тривалості обробки без попереднього детектора, мс	102	418	520	589	827
Математичне очікування тривалості обробки з попереднім детектором, мс	6	24	35	36	70
Дисперсія тривалості обробки з попереднім детектором, мс	3	4	5	6	13

Міра відгуку кута – безрозмірна величина, що оцінює ступінь схожості околиці точки на кут [5].

Для визначення зміни тривалості обробки кадру з використанням попереднього детектора було реалізовано тестове додаток, що замірює тривалість обробки кадру. Для кожного з представлених дозволів тестового зображення, наведеного рисунку 3.4, проводилося 1000 вимірів тривалості обробки, після чого на підставі отриманих даних розраховувалося математичне очікування тривалості обробки кадрів, подане в таблиці 3.1 [6].

На підставі даних з таблиці 3.1 був побудований графік залежності тривалості обробки зображення від кількості точок зображення, зображений рисунку 3.5. Отримана залежність тривалості виділення особливих точок від кількості пікселів для обраного обчислювача є лінійною та описується формулою

$$t(N) = 9 \times 10^{-5} N,$$

де  $N$  - Число пікселів .

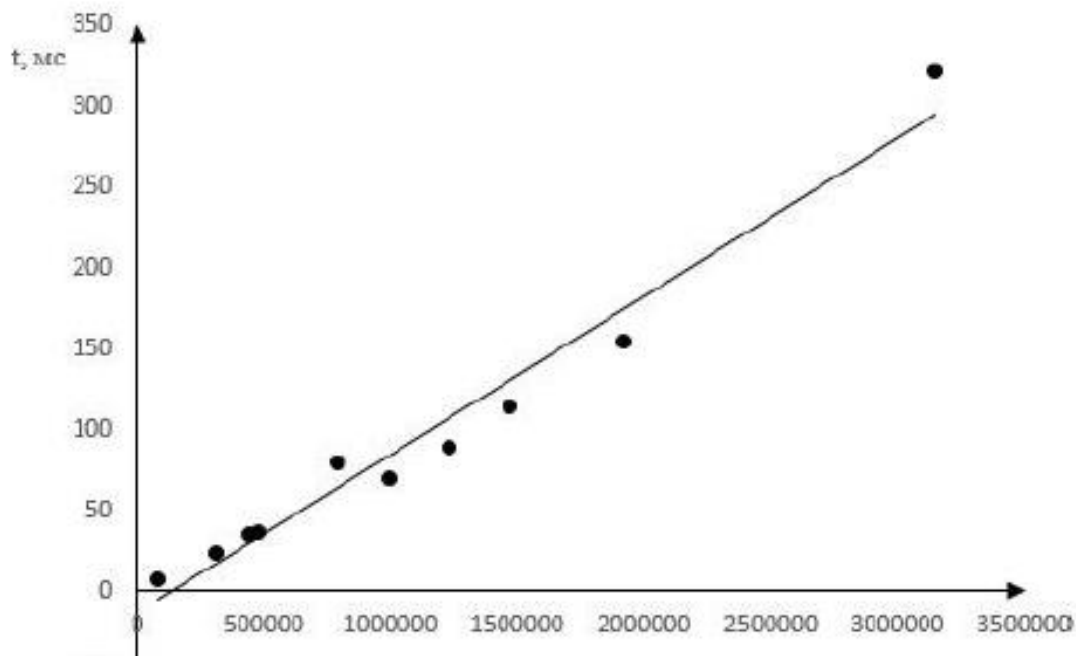


Рисунок 3.5 – Залежність тривалості обробки зображення від кількості точок

Як слід з таблиці 3.1, при детектуванні локальних особливостей тестового зображення детектором Харріса використання яскравого попереднього детектора дозволило скоротити тривалість детектування приблизно 220 раз.

Використання попереднього детектора суттєво впливає на кількість знайдених особливих точок. Детектор Харріса виділяє особливі точки на основі значень відгуку кута, тоді як попередній детектор використовує перепади яскравості зображення, які характерні не тільки для кутів, але і для прямих ліній, а також виділяються точок [5, 6].

Для виміру впливу попереднього детектора на кількість знайдених локальних особливостей зображення був зроблено підрахунок знайдених на тестовому зображенні особливих точок детектором Харріса як із використанням яскравого попереднього детектора, і без нього (рисунок 3.5). Результати вимірювань наведено у таблиці 3.2.

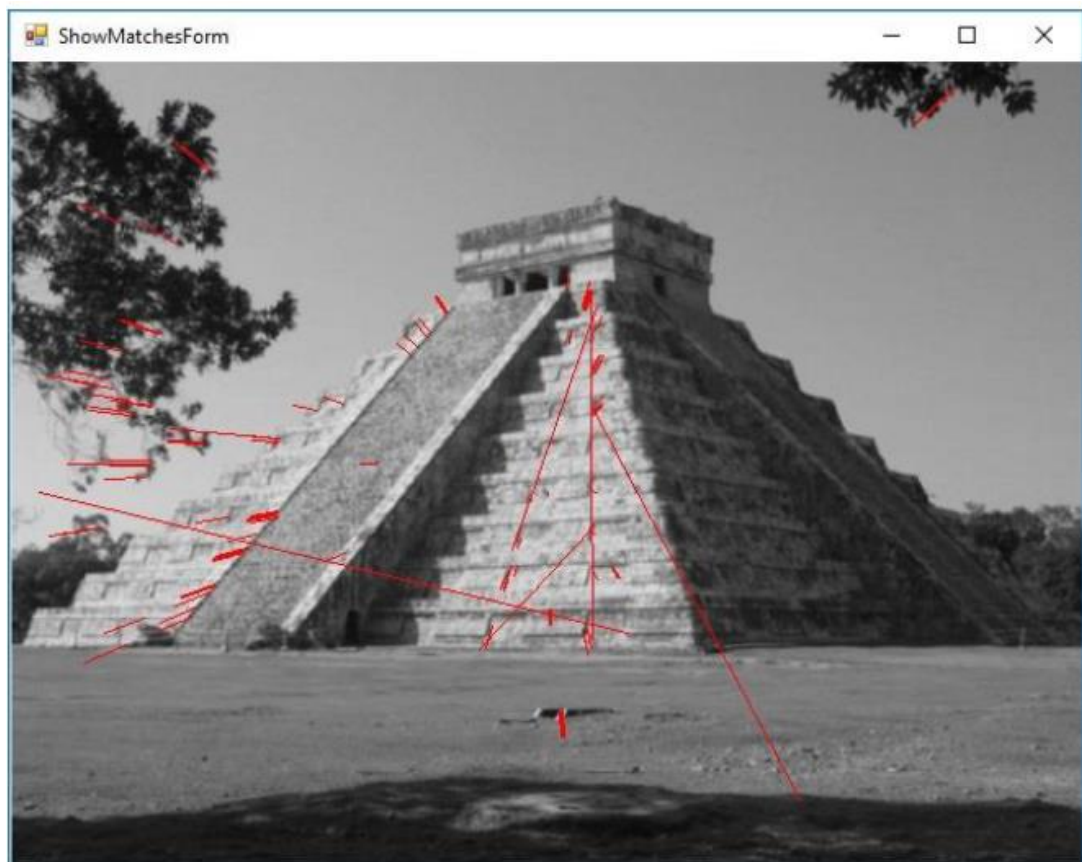


Рисунок 3.6– Тестове зображення зі зміненими особливими точками

Таблиця 3.2 – Кількість знайдених локальних особливостей

Роздільна здатність тестового зображення , пікселі	Кількість знайдених локальних особливостей із використанням попереднього детектора	Кількість знайдених локальних особливостей без використання попереднього детектора
320x240	8	54
640x480	5	173
768x576	4	264
800x600	14	133
1024x768	6	12

Як видно з таблиці 3.2, використання яскравого попереднього детектора значно скорочує кількість виділених особливостей зображення. Слід відзначити, що всі точки, виділені з використанням попереднього детектора, входять у безліч точок, які були виділені без його використання.

Причина скорочення кількості виділених локальних особливостей полягає в тому що попередній детектор відсікає кути з незначно вираженим градієнтом інтенсивності зображення, залишаючи тільки ті, які мають перепад яскравості вище граничного значення (3.3).

Вимірювання впливу використання яскравого попереднього детектора на відгук кута знайдених особливих точок проводилося нижчеописаним чином. Були взяті копії тестового зображення в різних дозволах, які були проаналізовані детектором Харріса з використанням яскравого попереднього детектора та без нього. В отриманих множин виділених особливостей зображення було виміряно максимальне та мінімальне значення відгуку кута, потім на основі цих вимірювань було розраховано порогове значення відгуку. Дані вимірювань наведено у таблиці 3.3.

Таблиця 3.3 – Значення відгуку кута виявлених особливих точок

Розмір тестового зображення, пікселі	320x240	640x480	768x576	800x600	1024x768
Мінімальне значення відгуку кута без ЯПД	$6,48 \cdot 10^{-5}$	$2,5 \cdot 10^{-5}$	$2,13 \cdot 10^{-5}$	$1,88 \cdot 10^{-5}$	$1,25 \cdot 10^{-5}$
максимальне значення відгуку кута без ЯПД	9,44	8,33	8,32	9,445	12,26
Порогове значення відгуку кута без ЯПД	6,68	5,89	5,89	6,67	8,67
Мінімальне значення відгуку кута з ЯПД	0,289	0,12	0,08	0,122	0,038
Максимальне значення відгуку кута з ЯПД	8,62	7,28	8,32	9,27	9,97

Зміна порогового значення відгуку кута пов'язано з тим, що точки, що мають максимальне значення відгуку кута не були виділено попереднім детектором, тому значення перепаду яскравості не перевищило порогове

значення. Таким чином, використання попереднього детектора суттєво скорочує безліч аналізованих точок зображення та кілька знижує їх якість.

В ході емуляційного обчислювального експерименту отримані дані, що свідчать про те, що використання яскравого попереднього детектора локальних особливостей зображення суттєво скорочує час виділення локальних особливостей зображення, що дозволяє використовувати детектори кутів у відеоінформаційних системах доповненої реальності, що працюють у реальному часу. Водночас під час використання попереднього детектора значно знижується кількість виділених локальних особливостей.

Для підвищення кількості виділених особливостей можливо застосування різних фільтрів шумів на етапі попереднього детектування, а також зниження порогового значення яскравості в попередньому детекторі за допомогою чисельних методів. Зниження заходи відгуку кута у виділених локальних особливостях пояснюється непопаданням деяких локальних особливостей у безліч точок, що виділяються попереднім детектором, що призводить до зміни порогового значення відгуку кута.

## ВИСНОВКИ

Проведено аналітичний огляд літератури, дано визначення терміну доповнена реальність, наведено класифікацію систем доповненої реальності. Також були виділені основні завдання, які вирішуються при реалізації таких систем та виділені основні складності.

Було зроблено порівняння найчастіше використовуваних детекторів і дескрипторів локальних особливостей зображення, виділені зі слабких боків і недоліки.

В ході обчислювального експерименту отримані дані, свідчать про те, що використання яскравого попереднього детектора локальних особливостей зображення суттєво скорочує час виділення локальних особливостей зображення, що дозволяє використовувати детектори кутів у відеоінформаційних системах доповненої реальності, що працюють у реальному часу. Водночас під час використання попереднього детектора значно знижується кількість виділених локальних особливостей.

## ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Зубенко В.П., Жовтоніжко Д.С., Можасєв О.О., Гук А.С. Дослідження детекторів локальних особливостей зображення. Системи управління, навігації та зв'язку. Полтава : Національний університет «Полтавська політехніка імені Юрія Кондратюка», 2024. Вип. 2(76). С. 66–70.
2. Ojala T. Multiresolution Gray Scale and Rotation Invariant Texture Classification with Local Binary Patterns / T. Ojala, M. Pietikäinen, T. Mäenpää // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2002. – № 7(24). – pp. 971987.
3. Zhao G. Dynamic Texture Recognition Using Local Binary Patterns with an Application to Facial Expressions / G. Zhao, M. Pietikainen // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2007. – № 6(29). – pp. 915-928.
4. Sun J. New Local Edge Binary Patterns for Image Retrieval / J. Sun, G. Fan, X. Wu // 20th IEEE International Conference on Image Processing. – 2013. – pp. 40144018.
5. Liao S. Dominant Local Binary Patterns for Texture Classification / S. Liao, M. W. K. Law, A. C. S. Chung // IEEE Transactions On Image Processing. – 2009. – №5 (18). – pp. 1107-1118.
6. Xie L. Video-based Facial Expression Recognition Using Histogram Sequence of Local Gabor Binary Patterns from Three Orthogonal Planes / L. Xie, H. Wei, W. Yang // Proceedings of the 33rd Chinese Control Conference. – 2014.
7. She R., Sun M. Security Energy Efficiency Analysis of CR-NOMA Enabled IoT Systems for Edge-cloud Environment. *Int. Journal of Computational Intelligence Systems*. 2023. Vol. 16(1), 118. DOI: <http://dx.doi.org/10.1007/s44196-023-00273-y>.
8. Kuchuk G., Nechausov S., Kharchenko, V. Two-stage optimization of resource allocation for hybrid cloud data store. *Int. Conf. on Information and*

*Digital Technologies*. Zilina, 2015. P. 266-271. DOI: <http://dx.doi.org/10.1109/DT.2015.7222982>.

9. Кучук Г.А., Коваленко А. А., Лукова-Чуйко Н. В. Метод мінімізації середньої затримки пакетів у віртуальних з'єднаннях мережі підтримки хмарного сервісу. *Системи управління, навігації та зв'язку*. Полтава . ПНТУ, 2017. Вип. 2(42). С. 117-120.

10. Sharma, M., Kaur, P. Reliable federated learning in a cloud-fog-IoT environment. *Journal of Supercomputing*. 2023. Vol. 79(14). P. 15435–15458. DOI: <http://dx.doi.org/10.1007/s11227-023-05252-w>.

11. Baucas, M.J., Spachos, P. Improving Remote Patient Monitoring Systems Using a Fog-Based IoT Platform with Speech Recognition. 2023. *IEEE Sensors Journal*. Vol. 23(15). P. 17611–17618. DOI: <http://dx.doi.org/10.1109/JSEN.2023.3287916>.

12. Number of Internet of Things (IoT) connected devices worldwide from 2019 to 2023, with forecasts from 2022 to 2030. Statista. Telecommunications. URL: <https://www.statista.com/statistics/1183457/iot-connected-devices-worldwide>.

13. Kowalczyk A. European IoT Spending to Reach Nearly \$227 Billion in 2023, Despite Ongoing Market Uncertainty, Says IDC. June 2023. URL: [https://www.idc.com/getdoc.jsp?containerId=prEUR250941023&utm\\_medium=embedd&utm\\_campaign=idc\\_embedd&utm\\_source=referral](https://www.idc.com/getdoc.jsp?containerId=prEUR250941023&utm_medium=embedd&utm_campaign=idc_embedd&utm_source=referral).