

БАКАЛАВРСЬКА РОБОТА

БР. ІІІ - 75.00.00.000 ІІЗ

Група ІІІ-22-3

Лялька Назар

2026

Івано-Франківський національний технічний університет нафти і газу

Факультет інформаційних технологій

Кафедра інженерії програмного забезпечення

Лялька Назар Іванович

(прізвище, ім'я, по батькові)

УДК 004.942

(індекс)

БАКАЛАВРСЬКА РОБОТА

Інтеграція комп'ютерного зору в мобільні додатки доповненої реальності

(назва роботи)

Інженерія програмного забезпечення

(назва освітньої програми)

121 - Інженерія програмного забезпечення

(шифр і назва спеціальності)

Робота містить результати власних досліджень, використання ідей, результатів і текстів інших авторів мають посилання на відповідне джерело:

Здобувач освітнього ступеня Лялька Н.І.
(підпис, ініціали та прізвище здобувача)

Науковий керівник Юрчишин Володимир Миколайович, д.т.н., професор
(підпис, прізвище, ім'я, по батькові, науковий ступінь, вчене звання керівника)

Допущено до захисту

Завідувач кафедри

доц. Бандура В.В.
(посада) (підпис) (дата) (ініціали та прізвище)

Івано-Франківський національний технічний університет нафти і газу
Інститут, факультет інформаційних технологій
Кафедра інженерії програмного забезпечення
Ступінь вищої освіти бакалавр
Спеціальність 121 – Інженерія програмного забезпечення

ЗАТВЕРДЖУЮ:

Зав. кафедрою _____ **ІПЗ**
доцент. _____ **В.В. Бандура**
“ _____ ” _____ 2026 р.

ЗАВДАННЯ
НА БАКАЛАВРСЬКУ РОБОТУ СТУДЕНТОВІ

Ляльці Назару Івановичу

(прізвище, ім'я, по-батькові)

1. Тема проекту (роботи) " Інтеграція комп'ютерного зору в мобільні додатки доповненої реальності "

керівник проекту (роботи) Юрчишин Володимир Миколайович, д.т.н., професор

затверджені наказом вищого навчального закладу від “ 21 ” квітня 2026 р. № 179/7

2. Строк подання студентом проекту (роботи) 10 червня 2026 р.

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

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

1. Аналіз вимог до програмного забезпечення

2. Аналіз вимог і постановка задачі

3. Проектування системи

4. Реалізація проекту

5. Тестування та впровадження

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

1 Приклад використання комп'ютерного зору (рис. 1.5, ст. 26)

2 Регулювання нахилу голови для WIM (рис. 2.3, ст. 31)

4. Монокулярні та бінокулярні наголовні дисплеї з прозорою оптикою (рис. 2.5, ст. 46)

4 Приклад комп'ютерного зору для AR-додатків (рис. 2.6, ст. 49)

5 Приклади об'єктів із виділеними елементами для використання в доповненої реальності. рис. 3.6, ст. 60)

6. Консультанти розділів проекту (роботи)

Розділ	Консультант	Підпис, дата	
		Завдання видав	Завдання прийняв

2. Дата видачі завдання 2026 р.

Керівник

_____ (підпис)

Завдання прийняв до виконання

_____ (підпис)

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів дипломного проекту (роботи)	Строк виконання етапів проекту	Примітка
1	Визначення та обґрунтування теми роботи	17.01.2026	виконано
2	Огляд існуючих концепцій, рішень та сервісів в даній області	21.02.2026	виконано
3	Побудова моделі або алгоритму власного рішення	01.03.2026	виконано
4	Документування реалізації власного оригінального рішення вибраними засобами	21.04.2026	виконано
5	Оформлення пояснювальної записки бакалаврської роботи	02.05.2026	виконано

Студент

_____ Лялька Н.І.
(підпис)

Керівник роботи

_____ Юрчишин В.М.
(підпис)

АНОТАЦІЯ

Бакалаврська робота містить 77 сторінок, 29 рисунків, 1 таблицю, список використаних джерел із 35 найменувань.

Метою роботи є дослідження принципів інтеграції комп'ютерного зору в мобільні додатки доповненої реальності та аналіз методів побудови ефективних, продуктивних і адаптивних AR-систем.

Об'єкт дослідження: мобільні додатки доповненої реальності.

Предмет дослідження: методи, алгоритми та інструменти комп'ютерного зору, що використовуються в мобільних AR-додатках.

Результати дослідження: проведено аналіз сучасних підходів до інтеграції комп'ютерного зору в системи доповненої реальності, досліджено архітектуру мобільних AR-додатків, розглянуто методи обробки зображень, відстеження об'єктів та розпізнавання сцен.

У першому розділі розглянуто теоретичні основи комп'ютерного зору та його роль у системах доповненої реальності, архітектуру мобільних AR-систем і основні проблеми їх реалізації.

У другому розділі досліджено методи комп'ютерного зору для мобільних AR-додатків, виконано порівняльний аналіз технологій доповненої та віртуальної реальності, а також розглянуто технічні вимоги до розробки.

У третьому розділі описано процес проєктування та реалізації мобільного AR-додатку з використанням комп'ютерного зору, а також проведено тестування й оцінку ефективності алгоритмів.

Висновок: інтеграція комп'ютерного зору в мобільні AR-додатки дозволяє значно розширити їх функціональні можливості, підвищити точність взаємодії з реальним середовищем і забезпечити більш природний користувацький досвід.

КЛЮЧОВІ СЛОВА: КОМП'ЮТЕРНИЙ ЗІР, ДОПОВНЕНА РЕАЛЬНІСТЬ, МОБІЛЬНІ ДОДАТКИ, ОБРОБКА ЗОБРАЖЕНЬ, ВІДСТЕЖЕННЯ ОБ'ЄКТІВ, AR, COMPUTER VISION.

ANNOTATION

The thesis consists of 75 pages, 28 figures, and a reference list with 35 sources.

The aim of the work is to study the principles of integrating computer vision into mobile augmented reality applications and to analyze methods for building efficient, high-performance, and adaptive AR systems.

Research object: mobile augmented reality applications.

Research subject: methods, algorithms, and tools of computer vision used for integration into mobile AR applications.

Research results: an analysis of modern approaches to integrating computer vision into augmented reality systems was conducted, the architecture of mobile AR applications was studied, image processing methods, object tracking, and scene recognition techniques were examined, and technical requirements for implementing such systems were defined.

The first section describes the theoretical foundations of computer vision and its role in augmented reality systems, the architecture of mobile AR applications, and the main challenges of their implementation.

The second section analyzes computer vision methods for mobile AR applications, provides a comparative analysis of augmented and virtual reality technologies, and reviews development tools and technical requirements.

The third section covers the design and implementation of a mobile AR application using computer vision, as well as testing and evaluation of tracking and recognition algorithm performance.

Conclusion: the integration of computer vision into mobile AR applications significantly enhances their functionality, improves interaction accuracy with the real environment, and provides a more natural user experience.

KEY WORDS: COMPUTER VISION, AUGMENTED REALITY, MOBILE APPLICATIONS, IMAGE PROCESSING, OBJECT TRACKING, AR, COMPUTER VISION.

ЗМІСТ

ВСТУП	8
РОЗДІЛ 1. ТЕОРЕТИЧНІ ОСНОВИ ІНТЕГРАЦІЇ КОМП'ЮТЕРНОГО ЗОРУ В МОБІЛЬНІ AR-ДОДАТКИ	10
1.1. Основи комп'ютерного зору та роль у системах доповненої реальності	10
1.2. Архітектура та принципи функціонування мобільних систем доповненої реальності	14
1.3. Основні проблеми та обмеження інтеграції комп'ютерного зору в AR-додатках	24
1.4 Висновок по розділу	27
РОЗДІЛ 2. МЕТОДИ ТА ІНСТРУМЕНТИ КОМП'ЮТЕРНОГО ЗОРУ В ДОПОВНЕНІЙ РЕАЛЬНОСТІ	28
2.1. Порівняльний аналіз технологій доповненої та віртуальної реальності	28
2.2. Методи комп'ютерного зору для мобільних AR-систем	32
2.3. Інструментальні засоби та технічні вимоги до реалізації AR-додатків з використанням комп'ютерного зору	38
2.4 Висновок по розділу	54
РОЗДІЛ 3. РОЗРОБКА ТА ІНТЕГРАЦІЯ КОМП'ЮТЕРНОГО ЗОРУ В МОБІЛЬНИЙ AR-ДОДАТОК	55
3.1. Проектування архітектури мобільного AR-додатку з використанням комп'ютерного зору	55
3.2. Програмна реалізація системи комп'ютерного зору в AR-додатку	59
3.3. Тестування та оцінка ефективності алгоритмів відстеження та розпізнавання	75
3.4 Висновок по розділу	76
ВИСНОВКИ	77
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ	79

					БР.ІП – 75.00.00.000 ПЗ						
Змн.	Арк.	№ докум.	Підпис	Дата							
Розроб.		Лялька Н.І.			Інтеграція комп'ютерного зору в мобільні додатки доповненої реальності			Літ.	Арк.	Акрушів	
Перевір.		Юрчишин В.М.								7	
Реценз.		Процюк Г.Я						ІФНТУНГ ІП-22-3			
Н. Контр.		Піх М.М.									
Затверд.		Бандура В. В.									
					Пояснювальна записка						

ВСТУП

У сучасному світі стрімкий розвиток мобільних технологій та зростання обчислювальних можливостей портативних пристроїв сприяють активному впровадженню інноваційних рішень у сфері доповненої реальності. Доповнена реальність (AR) стає важливим інструментом у таких галузях, як освіта, медицина, промисловість, навігація та розваги. Одним із ключових факторів ефективного функціонування AR-систем є використання методів комп'ютерного зору, що забезпечують аналіз зображень, розпізнавання об'єктів та їх точне відстеження у реальному середовищі. Інтеграція комп'ютерного зору в мобільні додатки доповненої реальності відкриває нові можливості для створення інтелектуальних, адаптивних і інтерактивних систем, однак водночас ставить перед розробниками низку складних технічних завдань.

Актуальність теми зумовлена необхідністю підвищення точності, продуктивності та надійності мобільних AR-додатків, що працюють у реальному часі та обмежених ресурсах мобільних пристроїв. Існуючі рішення часто стикаються з проблемами обробки великих обсягів візуальних даних, затримок у відстеженні об'єктів, нестабільності роботи в умовах змінного освітлення та складних сцен. У зв'язку з цим виникає потреба у дослідженні та впровадженні ефективних методів комп'ютерного зору, оптимізації алгоритмів і використанні сучасних інструментів для розробки мобільних AR-додатків. Розв'язання цих задач дозволить підвищити якість взаємодії користувача з цифровим контентом і розширити сфери застосування доповненої реальності.

Метою роботи є дослідження принципів інтеграції комп'ютерного зору в мобільні додатки доповненої реальності та розробка ефективного AR-додатку з використанням сучасних методів обробки зображень і відстеження об'єктів.

Завданнями дослідження є: аналіз предметної області та сучасних підходів до інтеграції комп'ютерного зору в AR-системи; дослідження архітектури мобільних додатків доповненої реальності; аналіз методів комп'ютерного зору, зокрема обробки зображень, розпізнавання та відстеження об'єктів; визначення функціональних і нефункціональних вимог до мобільного

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						9
Змн.	Арк.	№ докум.	Підпис	Дата		

AR-додатку; вибір інструментів та технологій для реалізації системи; проектування архітектури програмного забезпечення; реалізація мобільного додатку з інтеграцією комп'ютерного зору; тестування та оцінка ефективності розробленого рішення.

Об'єктом дослідження є мобільні додатки доповненої реальності та процеси їх розробки.

Предметом дослідження є методи, алгоритми та програмні засоби комп'ютерного зору, що застосовуються для інтеграції в мобільні AR-додатки.

Методи дослідження включають аналіз наукових джерел для вивчення теоретичних основ комп'ютерного зору та доповненої реальності, порівняльний аналіз існуючих технологій, моделювання архітектури програмної системи, експериментальні методи для реалізації та тестування програмного забезпечення, а також оцінювання ефективності алгоритмів.

Практичне значення отриманих результатів полягає у розробці мобільного AR-додатку з інтегрованими методами комп'ютерного зору, що забезпечує розпізнавання та відстеження об'єктів у реальному часі. Запропоноване рішення може бути використане в освітніх, навігаційних або інформаційних системах, а також слугувати основою для подальших досліджень у сфері доповненої реальності.

Структура роботи відповідає поставленій меті та завданням. У першому розділі розглянуто теоретичні основи комп'ютерного зору та доповненої реальності, а також проблеми їх інтеграції. У другому розділі досліджено методи та інструменти комп'ютерного зору, що використовуються в AR-додатках. У третьому розділі описано процес проектування, реалізації та оцінки ефективності мобільного AR-додатку.

Бакалаврська робота містить 75 сторінок, 28 рисунків, 1 таблицю, 3 розділи, список використаних джерел із 40 найменувань.

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						10
Змн.	Арк.	№ докум.	Підпис	Дата		

РОЗДІЛ 1 ТЕОРЕТИЧНІ ОСНОВИ ІНТЕГРАЦІЇ КОМП'ЮТЕРНОГО ЗОРУ В МОБІЛЬНІ AR-ДОДАТКИ

1.1 Основи комп'ютерного зору та його роль у системах доповненої реальності

Зі збільшенням потужності комп'ютерів та зменшенням їх розмірів, нові мобільні, носимі та повсюдні обчислювальні програми швидко стають реальними, надаючи людям доступ до онлайн-ресурсів завжди та всюди. Ця нова гнучкість робить можливими нові види програм, які використовують навколишній контекст людини. Доповнена реальність (AR) пропонує особливо потужний інтерфейс користувача (UI) для контекстно-залежних обчислювальних середовищ. Системи AR інтегрують віртуальну інформацію у фізичне середовище людини, щоб вона сприймала цю інформацію як таку, що існує в її оточенні. Мобільні системи доповненої реальності (MARS) надають цю послугу, не обмежуючи місцезнаходження людини спеціально обладнаною зоною. В ідеалі, вони працюють практично будь-де, додаючи відчутний шар інформації до будь-якого середовища, коли це необхідно. Роблячи це, вони мають потенціал для революції у способі представлення інформації людям. Матеріал, представлений комп'ютером, безпосередньо інтегрований з реальним світом, що оточує вільно пересувається людину, яка може взаємодіяти з ним, щоб відображати пов'язану інформацію, ставити та вирішувати запити, а також співпрацювати з іншими людьми. Світ стає інтерфейсом користувача.

Термін «доповнена реальність» (AR) використовується для опису комбінації технологій, що дозволяють у режимі реального часу змішувати контент, згенерований комп'ютером, з відображенням відео в реальному часі. AR базується на методах, розроблених у VR [1], і взаємодіє не лише з віртуальним світом, але й має певний ступінь взаємозалежності з реальним світом. Як зазначено, «доповнення» реальності саме по собі не має сенсу. Однак цей термін набуває сенсу, як тільки ми переорієнтовуємось на людину та її

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						11
Змн.	Арк.	№ докум.	Підпис	Дата		

сприйняття світу. Реальність не можна збільшити, але її сприйняття можна. Однак ми збережемо термін «доповнена реальність», навіть якщо розуміємо його як «збільшене сприйняття реальності».

Провели цінні та змістовні дослідження в галузі доповненої реальності. Однак останнє десятиліття було особливо багатим на досягнення в цій зростаючій галузі досліджень, що відкрило перспективи для кількох можливостей використання AR у різних сферах застосування. Наскільки нам відомо, у літературі немає оновлених досліджень, які б цілісним чином розглядали технології AR з огляду на численні сфери застосування, вплив мобільних технологій та взаємозв'язок між AR та віртуальною реальністю (VR). Для тих, хто хоче ознайомитися з галуззю AR, це дослідження надає огляд останніх технологій, потенційних застосувань, обмежень та майбутніх тенденцій систем AR.

Доповнена реальність пов'язана з концепцією *віртуальної реальності* (VR). VR намагається створити штучний світ, який людина може інтерактивно переживати та досліджувати, переважно через зір, але також за допомогою аудіо, тактильних та інших форм зворотного зв'язку. AR також забезпечує інтерактивний досвід, але прагне доповнити реальний світ, а не створювати повністю штучне середовище [2]. Фізичні об'єкти навколо людини Різні дослідники дотримуються вужчих або ширших визначень того, що саме являє собою AR (доповнена реальність). Хоча дослідницька спільнота здебільшого погоджується з більшістю елементів AR-систем, чому сприяють обміни досвідом та дискусії на кількох міжнародних конференціях у цій галузі, все ще існують невеликі розбіжності в думках та номенклатурі. Ми визначимо AR-систему як таку, що поєднує реальну та комп'ютерно-генеровану інформацію в реальному середовищі, інтерактивно та в режимі реального часу, та поєднує віртуальні об'єкти з фізичними. Водночас, AR є підгалуззю ширшої концепції *змішаної реальності* (MR), яка також включає симуляції, що переважно відбуваються у віртуальній сфері, а не в реальному світі.

Мобільна доповнена реальність (AR) застосовує цю концепцію у справді

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						12
Змн.	Арк.	№ докум.	Підпис	Дата		

мобільних умовах, тобто подалі від ретельно обумовлених середовищ дослідницьких лабораторій та спеціалізованих робочих зон.

Після короткого огляду історії систем доповненої реальності (AR) у наступному підрозділі ми розглянемо компоненти, необхідні для створення мобільного досвіду доповненої реальності (AR). Хоча AR потенційно може доповнити фізичне середовище інформацією, яку сприймають усі людські органи чуття, візуальні та слухові накладання наразі є найпоширенішими способами доповнення. У випадку візуальної AR, комп'ютерна графіка просторово реєструється та накладається на реальні об'єкти за допомогою технологій відображення та відстеження, які ми описуємо в цьому розділі.

Термін «доповнена реальність» вперше був введений у 1990 році, якого попросили вдосконалити дорогі схеми та пристрої для маркування, що використовуються для керівництва робітниками на заводському цеху[4]. Він запропонував замінити великі фанерні дошки, які містили індивідуально розроблені інструкції з підключення для кожного літака, на апарат, що кріпиться на голові, який відображає схеми конкретного літака за допомогою високотехнологічних окулярів та проектує їх на багатоцільові дошки багаторазового використання [39].

Багато авторів погоджуються визначати AR таким чином, що це вимагає використання головних дисплеїв (HMD) [5]. Однак, щоб уникнути обмеження AR конкретними технологіями, ми пропонуємо визначати AR як системи, що мають такі характеристики:

- 1) поєднують реальне та віртуальне;
- 2) інтерактивні в реальному часі;
- 3) зареєстровані в 3D.

Це визначення має на меті дозволити іншим технологіям, таким як мобільні технології, окрім

Шлеми для моніторингу (HMD), зберігаючи при цьому основні компоненти доповненої реальності (AR) [6]. Двовимірні віртуальні накладання поверх живого відео можуть бути виконані з інтерактивною швидкістю, але ці

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						13
Змн.	Арк.	№ докум.	Підпис	Дата		

накладання не поєднуються з реальним світом у 3D [7]. Однак це визначення дозволяє використовувати інтерфейси на основі моніторів, монокулярні системи, прозорі HMD або мобільні пристрої.

Системи доповненої реальності побудовані на трьох основних елементах: відстеженні та реєстрації, технології відображення та рендерингу в реальному часі. По-перше, доповнена реальність – це технологія, яка повинна бути інтерактивною в реальному часі та зареєстрованою у трьох вимірах. При спробі досягти правдоподібного доповненого зображення важливе точне відстеження та реєстрація, оскільки для передачі правдоподібного зображення користувачеві реальна камера повинна бути зіставлена з віртуальною таким чином, щоб перспективи обох середовищ точно збігалися [8]. Особливо для користувача, що рухається, система повинна постійно визначати положення в середовищі користувача навколо віртуального об'єкта, оскільки згенерований комп'ютером об'єкт повинен виглядати нерухомим [8]. Якщо потрібна така форма повного відстеження з глобальною системою координат, можна розрізнити відстеження зовні всередину та зсередини назовні. Перший стосується систем, де в навколишньому середовищі розміщені датчики, які відстежують випромінювачі на мобільних об'єктах: наприклад, використання датчиків на основі глобальної системи позиціонування (GPS) для відстеження розташування мобільного пристрою або тріангуляції положення мобільного пристрою між телефонними щоглами [3]. Другий тип використовує внутрішні датчики, закріплені на мобільних об'єктах: камеру для візуального відстеження, цифровий компас для відстеження напрямку телефону, акселерометр для відстеження прискорення. Однак обидві ці системи мають свої недоліки, оскільки GPS, наприклад, не такий точний усередині будівель, як зовні, а візуальне відстеження сильно залежить від умов освітлення та видимості [10].

Розглядають як технологію відображення, так і рендеринг у реальному часі як основні будівельні блоки та виклики майбутнього. Перший пов'язаний з обмеженими оптичними, технічними та людськими факторами обмеженнями. Другий, рендеринг у реальному часі, пов'язаний зі здатністю пристроїв

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						14
Змн.	Арк.	№ докум.	Підпис	Дата		

доповненої реальності розміщувати шар графічних елементів поверх реального середовища швидко та реалістично. Кінцевою метою, є інтеграція згенерованого комп'ютером об'єкта таким чином, щоб користувач не міг розрізнити реальне та віртуальне.

1.2 Архітектура та принципи функціонування мобільних систем доповненої реальності

Переглядаючи наше визначення AR, ми можемо визначити компоненти, необхідні для MARS. Для початку потрібна *обчислювальна платформа*, яка може генерувати та керувати віртуальним матеріалом, що накладається поверх фізичного середовища, обробляти інформацію трекера та керувати AR-дисплеєм(ями).

Далі, потрібні *дисплеї* для представлення віртуального матеріалу в контексті фізичного світу [4]. У випадку доповнення візуального сприйняття це можуть бути дисплеї, що носяться на голові, мобільні портативні дисплеї або дисплеї, інтегровані у фізичний світ. Інші органи чуття (слух, дотик або нюх) також можуть бути потенційно доповнені. Зокрема, просторовий звук часто використовується для передачі локалізованої інформації, доповнюючи або повністю замінюючи візуальні елементи.

реєстрацію: вирівнювання віртуальних елементів з фізичними об'єктами, які вони анотують. Для візуальної та слухової реєстрації це можна зробити, *відстежуючи* положення та орієнтацію голови користувача та пов'язуючи це вимірювання з моделлю середовища та/або змушуючи комп'ютер «бачити» та потенційно інтерпретувати середовище за допомогою камер та комп'ютерного зору.

Носимі технології введення та взаємодії дозволяють мобільній людині працювати з доповненим світом (наприклад, робити вибір або отримувати доступ до баз даних, що містять відповідний матеріал, та візуалізувати їх) та додатково доповнювати навколишній світ. Вони також дають змогу людині

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						15
Змн.	Арк.	№ докум.	Підпис	Дата		

спілкуватися та співпрацювати з іншими користувачами MARS.

Бездротова мережа необхідна для зв'язку з іншими людьми та комп'ютерами під час руху. Динамічна та гнучка мобільна доповнена реальність (AR) спиратиметься на інформацію, що зберігається до секунди, яку неможливо зберегти на обчислювальному пристрої до запуску програми. Наприклад, це дозволить повідомляти зайнятому пасажирові про затримки поїздів чи автобусів та дорожній рух.

Це підводить нас до ще одного пункту у списку вимог до MARS: *технології зберігання та доступу до даних*. Якщо MARS має надавати інформацію про поточне середовище роумінгу, йому потрібно звідкись отримувати дані про це середовище. Сховища даних повинні надавати інформацію, що відповідає поточному контексту роумінгу. Виявлення, управління та доступ до даних і послуг ставлять кілька дослідницьких питань, які вивчаються дослідниками у спільнотах баз даних, проміжного програмного забезпечення та контекстно-орієнтованих послуг. З точки зору користувача, важливим питанням є те, як отримати найрелевантнішу інформацію з найменшими зусиллями та як мінімізувати інформаційне перевантаження.

Доповнена реальність (AR) – це надзвичайно перспективна та дедалі більш доцільна технологія інтерфейсу користувача, але поточні системи все ще є здебільшого дослідницькими прототипами. Немає очевидного найкращого рішення для багатьох складних областей у сфері доповненої реальності, таких як технології відстеження та відображення [5]. Нові результати досліджень постійно відкривають нові шляхи дослідження. Розробник, який планує розгорнути технологію доповненої реальності для конкретного завдання, повинен приймати проектні рішення, які оптимізують продуктивність доповненої реальності для даного застосування, на основі ретельного аналізу завдання. У деяких випадках можливо розробити спеціалізовані рішення доповненої реальності, які добре працюватимуть в обмежених зонах зі спеціальною апаратною підтримкою. З іншого боку, якщо аналіз завдання для сценарію доповненої реальності покаже, що кінцевий користувач системи

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						16
Змн.	Арк.	№ докум.	Підпис	Дата		

повинен отримувати підтримку в широкому спектрі місць та ситуацій, можливо, включаючи активний відпочинок на свіжому повітрі, тоді ми вступаємо у сферу справжньої мобільної доповненої реальності (AR).

Значний відсоток матеріалів конференцій, присвячених доповненій реальності (AR), надходить від промислових дослідницьких лабораторій або є спільною роботою університетів та промисловості. Багато ранніх публікацій з доповненої реальності є «прикладними роботами», що описують застосування нових технології в різних галузях. У решті цього розділу ми наводимо огляд потенційних застосувань мобільних систем доповненої реальності.

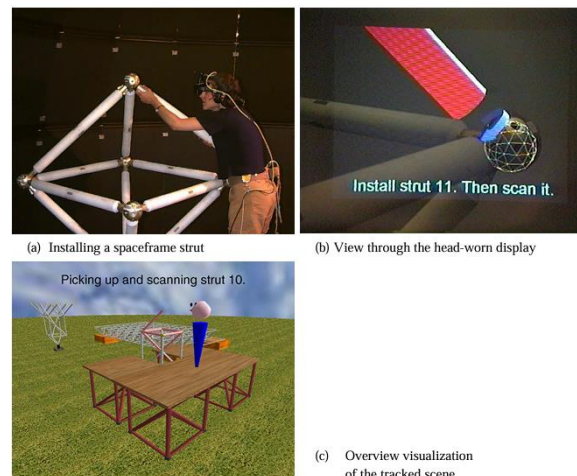


Рисунок 1.1 - AR для будівництва.

Збірка та конструкція. Протягом дев'яти років дослідники Boeing створили кілька ітерацій прототипів для складання електричних проводів для літаків за допомогою доповненої реальності (AR). AR накладала схематичні діаграми, а також супровідну документацію безпосередньо на дерев'яні дошки, на яких кабелі прокладалися, зв'язувалися та обгорталися. Комп'ютер вів (і потенційно міг розмовляти) складальників через процес монтажу. Оскільки отримані проводові пучки були достатньо довгими, щоб проходити через значні частини літака, стаціонарних AR-рішень було недостатньо, і проект перетворився на вправу з використання мобільної AR для конкретного сценарію застосування.

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						17
Змн.	Арк.	№ докум.	Підпис	Дата		

«Доповнена реальність для будівництва» – це прототип для будівництва каркасних конструкцій у приміщенні. Як показано на рисунку 1.1 (аб), будівельник бачитиме та чутиме через дисплеї на голові, де має бути встановлено наступний конструктивний елемент [6]. Будівельник сканує позначений елемент за допомогою зчитувача штрих-кодів до та після встановлення, щоб переконатися, що правильний елемент встановлено у правильному місці. Можливість створення віртуальних оглядових візуалізацій усієї будівельної сцени, як показано на прикладі малого масштабу за допомогою мережевої графіки, що оновлюється в режимі реального часу, в одній з демонстрацій прототипу (див. рисунок 1.1 (с)), є додатковою перевагою відстеження кожного окремого робітника та його дій, і може виявитися надзвичайно корисною для управління великими складними будівельними проектами. Такі будівельні завдання зрештою відбуватимуться на відкритому повітрі.



Рисунок 1.2 - Навігаційні інтерфейси доповненої реальності, зображення яких відображаються через головні дисплеї.

(а) Керування в приміщенні з використанням оглядової візуалізації та стрілок. (б) Віртуальні стежки та прапори на вулиці (вид з даху будівлі).

Технічне обслуговування та інспекція. Окрім складання та будівництва, інспекція та технічне обслуговування – це інші сфери виробництва, які можуть отримати значні переваги від застосування технологій MARS. Пропонують прототип системи доповненої реальності (AR) для інспекції електронних деталей

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						18
Змн.	Арк.	№ докум.	Підпис	Дата		

у межах виробничого підприємства з великою площею. Їхній рюкзак MARS відстежується за допомогою виключно інерційної системи відстеження (трекер орієнтації гіроскопа плюс датчик прискорення), а калібрування часто доводиться налаштовувати вручну. Пропонують використовувати візуальні кодовані маркери для великих промислових середовищ. Представляють системну архітектуру прототипу MARS для використання в технічному обслуговуванні атомних електростанцій. AR добре підходить для ситуацій, які потребують «рентгенівського зору», здатності бачити крізь тверді конструкції. Використовуючи прямі накладення прихованої інфраструктури, AR може допомогти працівникам з технічного обслуговування, які намагаються знайти обірване кабельне з'єднання в стінах будівлі або місцезнаходження протікання труби під поверхнею дороги [7]. Точне положення могло бути визначено автоматично (наприклад, встановленими датчиками), і в цьому випадку пряма візуалізація проблемної ділянки за допомогою доповненої реальності (AR) може бути найшвидшим способом спрямувати увагу працівника на потрібну область. Як альтернатива, AR може бути використана як допоміжний інструмент для визначення проблеми, оскільки вона може миттєво та безпосередньо візуалізувати будь-які дані, які працівник може зібрати, досліджуючи середовище за допомогою різних датчиків.

Навігація та пошук шляху. Враховуючи деякі інші способи використання мобільної доповненої реальності (AR) на свіжому повітрі, важливою сферою застосування носимих систем є їх використання як навігаційних засобів. Носимі комп'ютери можуть значно допомогти незрячим користувачам за допомогою аудіо та тактильного зворотного зв'язку. Якщо слухова інформація, що стосується реальних точок маршруту та особливостей навколишнього середовища, представлена користувачеві системи з відстеженням положення через просторово-утворений стереозвук, це чітко відповідає нашому визначенню AR-системи. Візуальна AR може допомогти в навігації, безпосередньо вказуючи місця в полі зору користувача за допомогою анотацій напрямку, таких як стрілки та стежки для слідування (див. рис. 1.2), або вказуючи на закриту

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						19
Змн.	Арк.	№ докум.	Підпис	Дата		

інфраструктуру, або безпосередньо за допомогою візуально кодованого зображення [8].

Туризм. Розвиток навігаційних ULS шляхом включення інформації про об'єкти в середовищі мобільної людини, які можуть бути цікавими для мандрівника, природним чином призводить до застосувань у туризмі. У цьому випадку AR використовується не лише для пошуку місць призначення, але й для відображення довідкової інформації. Наприклад, замість того, щоб шукати опис та історичну історію відомого собору в путівнику (або навіть на бездротово підключеному кишеньковому комп'ютері перед місцем розташування), AR може оживити повітря навколо церкви інформацією: 3D-моделі пов'язаного мистецтва чи архітектури, життя та творчість архітектора або архітектурні зміни протягом століть можуть бути задокументовані на місці за допомогою накладень. Можливості безмежні та обмежуються лише обсягом та типом інформації, доступної людині з підтримкою AR, та можливостями пристрою AR, який людина носить.

На рисунку 1.1 показано приклад мобільного доповненого путівника по ресторанах. Цей прототип MARS надає інтерфейс до бази даних ресторанів у Морнінгсайд-Хайтс, Нью-Йорк. Інформація про ресторани надається або у вигляді оглядової 3D-карти, щоб користувач міг дістатися до певного місця на свій вибір, або у вигляді прямих анотацій до самих місць розташування ресторанів. Вибравши заклад, користувач може відкрити спливаюче вікно з додатковою інформацією про нього: короткий опис, адреса та номер телефону, зображення інтер'єру та, доступне одним клацанням миші: меню та, якщо доступно, відгуки про ресторан, а також його зовнішня веб-сторінка. Ми обговоримо питання відображення пов'язаної інформації.

Географічні польові дослідження. Польові фахівці з географії та регіональних наук могли б використовувати методи доповненої реальності (AR) для збору, порівняння та оновлення даних опитувань і статистики в польових умовах. Допмагаючи збирати та відображати дані, система доповненої реальності могла б дозволити виявляти закономірності в польових умовах, а не

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						20
Змн.	Арк.	№ докум.	Підпис	Дата		

лише в лабораторії [9]. Було б можливою миттєва перевірка та порівняння інформації з даними у файлі.

Журналістика. Журналістика – це ще одна сфера, на яку MARS може мати значний вплив. Обговорює використання бездротових технологій для мобільного журналіста, який висвітлює та документує розвиток новин у русі. MARS можна використовувати для того, щоб залишати нотатки на місці події для інших журналістів та фотографів, що співпрацюють, для перегляду та реагування. Проект «Ситуаційні документальні фільми» в Колумбійському університеті, є результатом співпраці між інформатикою та журналістикою та використовує MARS для розповіді історій та представлення історичної інформації.

Архітектура та археологія. Доповнена реальність (AR) також особливо корисна для візуалізації невидимого: проєктів архітекторів мостів або будівель, які ось-ось будуть побудовані на певному місці; історичних будівель, давно знесених на їхньому первісному місці; або реконструкцій археологічних пам'яток. На рисунку 1.3, знятому через прозорий дисплей, що носить на голові, що представляє ситуаційний документальний фільм про історію Колумбії, показано модель частини психіатричної лікарні Блумінгдейл, яка колись займала головний кампус. Європейський спонсорований проєкт ARCHEOGUIDE має на меті реконструювати об'єкт культурної спадщини в доповненій реальності та дозволити відвідувачам оглянути та дізнатися про стародавню архітектуру та звичаї. Першим місцем, обраним для проведення випробувань, є стародавнє місто Олімпія в Греції.

(a) Користувач з рюкзаком MARS. (b) Вид крізь дисплей на голові. Віртуальні прапорці, що позначають цікаві місця. (c) Додатковий портативний інтерфейс для взаємодії з віртуальним матеріалом.

Моделювання міст. Доповнена реальність (AR) не використовуватиметься виключно для пасивного перегляду чи отримання інформації за допомогою випадкових кліків кнопок миші (або, точніше, гудзиків на сорочці). Багато дослідників досліджують, як технології AR можна

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						21
Змн.	Арк.	№ докум.	Підпис	Дата		

використовувати для введення інформації в комп'ютер. Одним із практичних прикладів є 3D-моделювання зовнішніх сцен: використання мобільної платформи для створення 3D-рендерів будівель та інших об'єктів, що моделюють саме середовище, щоб пізніше використовувати їх як фон для презентацій AR.

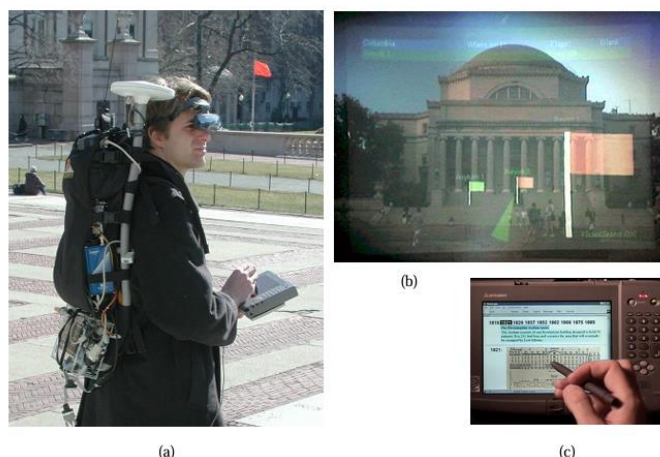


Рисунок 1.3 - Документальні фільми з витримкою подій.

Розваги. Додаток для ситуативних документальних фільмів також свідчить про потенціал технології для розважальних цілей. Замість того, щоб показувати 3D-фільми-«аттракціони», такі як популярна презентація «Термінатора 2» в Universal Studios, глядачам у спеціальних театрах тематичних парків, віртуальні актори у сценах зі спецефектами одного дня можуть заповнити самі вулиці тематичних парків, залучаючи гостей у костюмах доповненої реальності до видовищних дій. Як ранній початок у цьому напрямку, кілька дослідників експериментували із застосуванням мобільних технологій доповненої реальності в іграх.

Медицина. Доповнена реальність має важливі можливості застосування в медицині. Багато з них, такі як системи підтримки хірургів, які допомагають хірургам під час їхніх операцій за допомогою живих накладок, вимагають дуже точної реєстрації, але не вимагають, щоб хірург був надзвичайно мобільним під час підтримки системою доповненої реальності. Однак існує кілька можливих застосувань мобільної доповненої реальності в медичній галузі. У лікарні чи

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						22
Змн.	Арк.	№ докум.	Підпис	Дата		

будинку для людей похилого віку лікарі чи медсестри під час своїх візитів до пацієнтів можуть отримати важливу інформацію про кожного з них стан пацієнта безпосередньо передається на їхні окуляри. У польових умовах персонал швидкої медичної допомоги міг би швидше оцінити ситуацію за допомогою носимих датчиків та технології доповненої реальності (AR). Вони могли б прикріпити носимі датчики до пацієнта та з того часу могли б перевіряти стан пацієнта через окуляри доповненої реальності буквально одним поглядом. Крім того, віддалений експерт у віддаленій лікарні міг би бути залучений до процесу та спілкуватися з польовим працівником через систему доповненої реальності, бачачи через зображення з камер те, що бачить польовий працівник, що може бути важливим для підготовки до неминучої операції в лікарні.

Моніторинг інформації про здоров'я групи людей одночасно може бути корисним для тренерів або наставників під час спортивних тренувань або змагань. Військові також мають потенційне медичне застосування мобільних технологій доповненої реальності. Стан здоров'я солдатів на полі бою можна буде контролювати, і у разі будь-яких травм командир зможе отримувати візуалізацію місцезнаходження та стану поранених у режимі реального часу.

Військова підготовка та бойові дії. Військові дослідження призвели до розробки систем супутникової навігації та проєкційних дисплеїв для бойових пілотів. Військові дослідницькі лабораторії також вже деякий час досліджують потенціал мобільної технології доповненої реальності (AR) для наземних воїнів. Що стосується можливого використання AR у військових операціях, то існує значний збіг з цивільними застосуваннями на загальному рівні. Навігаційна підтримка, покращення зв'язку, ремонт та технічне обслуговування, а також невідкладна медицина є важливими темами як у цивільному, так і в військовому житті. Однак існують конкретні переваги, які технологія AR може принести військовому користувачеві. Більшість місій відбуваються на незнайомих територіях. Види карт, що природно проєктуються в обмежений огляд воїна на сцену бою, можуть надати додаткову інформацію про місцевість, яку нелегко помітити. Крім того, дані розвідки та інформація про планування місій можуть

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						23
Змн.	Арк.	№ докум.	Підпис	Дата		

бути інтегровані в ці інформаційні дисплеї, що уточнює ситуацію та окреслює конкретні підзавдання для окремих військ. Поточні дослідження у Військово-морській дослідницькій лабораторії спрямовані на те, як такі інформаційні дисплеї можуть бути найефективніше доставлені воїнам. Окрім використання в бойових умовах, мобільна доповнена реальність (AR) також може виявитися цінним військовим інструментом для навчання та симуляції. Наприклад, масштабні бойові сценарії можна було б випробувати з імітацією дій противника в реальних навчальних умовах.

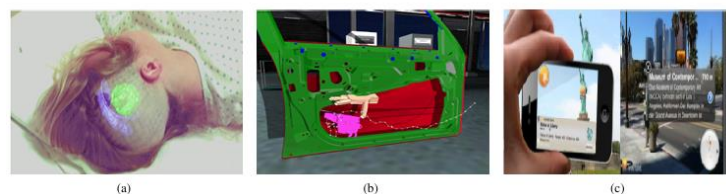


Рисунок 1.4 - Застосування доповненої реальності в (а) керованій хірургії; (б) складанні продукції; та (в) навігації в міських умовах.

Управління персональною інформацією та маркетинг. Можна лише здогадуватися, які зусилля в галузі мобільної доповненої реальності (AR) можуть зрештою призвести до комерційного успіху. Деякі невеликі компанії вже пропонують спеціалізовані AR-рішення (TriSense). Що б не призвело до широкого використання, найбільшим потенційним ринком для цієї технології можуть стати персональні носимі комп'ютери. AR може служити передовим та миттєвим інтерфейсом користувача для носимих комп'ютерів. У особистому щоденному використанні AR може підтримувати та інтегрувати поширені завдання, такі як спілкування електронною поштою та телефоном, з накладаннями, що враховують місцезнаходження, надавати навігаційні вказівки, дозволяти людям зберігати персональну інформацію, пов'язану з певними місцями, та забезпечувати єдиний інтерфейс керування для всіх видів побутової техніки вдома. Звичайно, така персональна платформа була б дуже привабливою для агентств прямого маркетингу. Магазины могли б пропонувати віртуальні

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						24
Змн.	Арк.	№ докум.	Підпис	Дата		

купони на знижку пішоходам, що проходять повз. Віртуальні білборди могли б рекламувати товари на основі профілю користувача. Віртуальні 3D-прототипи продуктів могли б з'являтися в окулярах клієнта. Щоб захистити особу від небажаної інформації, AR-платформа повинна включати відповідні механізми фільтрації та управління переглядами.

1.3 Основні проблеми та обмеження інтеграції комп'ютерного зору в AR-додатках

Незважаючи на великий потенціал мобільної доповненої реальності (AR) у багатьох сферах застосування, прогрес у цій галузі досі демонструється майже виключно за допомогою дослідницьких прототипів. Очевидно, що час для комерціалізації ще не зовсім настав. Запитуючи про причини цього, потрібно уважно подивитися на масштаб завдання. Хоча впроваджуються все кращі рішення для технічних проблем, пов'язаних з носимими комп'ютерами, залишається кілька проблемних областей, таких як мініатюризація технології введення/виведення, джерел живлення та тепловіддача, особливо в невеликих високопродуктивних системах. Також потрібна міцність. У деяких ранніх носимих системах фраза «знос», здавалося, досить влучно вказувала на жахливі наслідки використання. На додаток до цих стандартних вимог до носимих комп'ютерів, мобільна AR додає багато іншого: надійне та повсюдне відстеження положення в широкій зоні, точне та самокалібруюче відстеження орієнтації (голови); надлегкі, надяскраві, надпрозорі окуляри з широким полем зору; швидкі можливості 3D-графіки – це лише найважливіші вимоги. Ми розглянемо сучасні технології, що вирішують ці проблемні області, у наступному розділі. Якби для кожного з необхідних компонентів було обрано найкращі технології, доступні сьогодні, можна було б побудувати потужний (хоча й великий) MARS [10].

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						25
Змн.	Арк.	№ докум.	Підпис	Дата		



Рисунок 1.5 – Приклад використання комп'ютерного зору

Використання доповненої реальності на відкритому повітрі є особливо складним завданням, оскільки існує широкий спектр робочих умов, яким може піддаватися система. Більше того, на відміну від контрольованого середовища в приміщенні, на зовнішні умови мало впливу; наприклад, освітлення може варіюватися від прямих сонячних променів, можливо, посилених відбиваючим середовищем (наприклад, снігом), до абсолютної темряви без штучних джерел світла вночі. Зовнішні системи повинні витримувати всі можливі погодні умови, включаючи вітер, дощ, мороз та спеку [11].

Перелік викликів не обмежується технологіями на стороні користувача. Залежно від технології відстеження, системи доповненої реальності (AR) повинні мати доступ до моделі середовища, яке вони повинні анотувати, або вимагати, щоб таке середовище було *підготовлене* (наприклад, оснащене візуальними маркерами або електронними мітками). Відстеження на основі зору в непідготовленому середовищі наразі не є життєздатним загальним рішенням, але дослідження в цій галузі намагаються створити рішення для майбутніх систем. Дані, які будуть представлені в AR-накладаннях, повинні бути пов'язані з місцезнаходженням у середовищі. Повинен бути встановлений стандартний метод доступу для отримання таких даних з баз даних, що відповідають за область, через яку наразі проходить користувач MARS. Це вимагає таких механізмів, як автоматичне виявлення сервісів та визначення стандартних форматів обміну, які підтримують як сервери баз даних, так і програмне

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						26
Змн.	Арк.	№ докум.	Підпис	Дата		

забезпечення MARS. З історії стандартів протоколів зрозуміло, що без великого попиту та можливостей заробітку на горизонті прогрес у цих напрямках може бути повільним. З іншого боку, Всесвітня павутина, HTML та HTTP розвивалися з подібних початкових умов. Деякі дослідники розглядають геолокаційно-залежні обчислення в глобальному масштабі як законного наступника Всесвітньої павутини, якою ми її знаємо сьогодні.

У кіноіндустрії спецефекти бездоганно поєднують комп'ютерно згенеровані світи з реальними сценами. Наразі ці зусилля потребують днів і місяців рендерингу та дуже ретельної інтеграції віртуального матеріалу в реальні кадри. В доповненій реальності (AR) рендеринг не тільки має виконуватися в режимі реального часу, але й рішення про те, що відображати, та генерацію цього матеріалу мають ініціюватися та контролюватися на льоту. Зробити візуальні ефекти максимально інформативними, а в деяких випадках також максимально реалістичними, відтвореними з правильним освітленням для забезпечення безперебійного враження, є відкритим завданням для візуальної доповненої реальності.

1.4 Висновок по розділу

У першому розділі було розглянуто теоретичні основи інтеграції комп'ютерного зору в мобільні додатки доповненої реальності. Проаналізовано основи комп'ютерного зору та визначено його роль у забезпеченні функціональності AR-систем. Досліджено архітектуру та принципи функціонування мобільних систем доповненої реальності, що дозволяють поєднувати цифрові об'єкти з реальним середовищем. Також розглянуто основні проблеми та обмеження інтеграції комп'ютерного зору в AR-додатках, зокрема питання продуктивності, точності розпізнавання та обмежень мобільних пристроїв. Отримані результати формують теоретичну основу для подальшого дослідження методів та інструментів комп'ютерного зору в AR-системах.

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						27
Змн.	Арк.	№ докум.	Підпис	Дата		

РОЗДІЛ 2 . МЕТОДИ ТА ІНСТРУМЕНТИ КОМП'ЮТЕРНОГО ЗОРУ В ДОПОВНЕНІЙ РЕАЛЬНОСТІ

2.1 Порівняльний аналіз технологій доповненої та віртуальної реальності

Термін «віртуальна реальність» зазвичай використовується популярними ЗМІ для опису уявних світів, які існують лише в комп'ютерах та нашій свідомості. Однак давайте точніше визначимо цей термін. Віртуальне визначається як буття за суттю або ефектом, але не фактично. Реальність визначається як щось, що становить реальну або актуальну річ, на відміну від чогось, що є лише видимим; щось, що існує незалежно від ідей, що його уявляють. На щастя, [12] нещодавно визначино повний термін «віртуальна реальність» як штучне середовище, яке сприймається через сенсорні стимули (як зображення та звуки), що надаються комп'ютером, і в якому дії людини частково визначають те, що відбувається в середовищі. Також визначається віртуальну реальність як комп'ютерне середовище, з яким можна взаємодіяти так, ніби це середовище реальне. Гарна система віртуальної реальності дозволить користувачам фізично ходити навколо об'єктів і торкатися їх, ніби вони реальні. Іван Сазерленд, творець однієї з перших у світі систем віртуальної реальності, заявив: «Кінцевим експонатом, звичайно, була б кімната, в якій комп'ютер міг би контролювати існування матерії. Стілець, виставлений у такій кімнаті, був би достатньо гарним, щоб на ньому сидіти. Кайданки, виставлені в такій кімнаті, обмежували б, а куля, виставлена в такій кімнаті, була б смертельною».

Мобільна доповнена реальність

Зі збільшенням потужності комп'ютерів та зменшенням їх розмірів, нові мобільні, носимі та повсюдні обчислювальні програми швидко стають реальними, надаючи людям доступ до онлайн-ресурсів завжди та скрізь [13]. Ця нова гнучкість робить можливим новий клас програм, які використовують навколишній контекст людини [15]. Доповнена реальність вже пропонує

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						28
Змн.	Арк.	№ докум.	Підпис	Дата		

особливо потужний інтерфейс користувача (UI) для контекстно-залежних обчислювальних середовищ. Системи доповненої реальності інтегрують віртуальну інформацію у фізичне середовище людини, щоб вона сприймала цю інформацію як таку, що існує в її оточенні [16]. Мобільні системи доповненої реальності надають цю послугу, не обмежуючи місцезнаходження людини спеціально обладнаною зоною [9]. В ідеалі вони працюють віртуально.



Рисунок 2.1 - Адаптована схема континууму віртуальності.

Мобільна доповнена реальність

Зі збільшенням потужності комп'ютерів та зменшенням їх розмірів, нові мобільні, носимі та повсюдні обчислювальні програми швидко стають реальними, надаючи людям доступ до онлайн-ресурсів завжди та скрізь [14]. Ця нова гнучкість робить можливим новий клас програм, які використовують навколишній контекст людини [15]. Доповнена реальність вже пропонує особливо потужний інтерфейс користувача (UI) для контекстно-залежних обчислювальних середовищ. Системи доповненої реальності інтегрують віртуальну інформацію у фізичне середовище людини, щоб вона сприймала цю інформацію як таку, що існує в її оточенні [16]. Мобільні системи доповненої реальності надають цю послугу, не обмежуючи місцезнаходження людини спеціально обладнаною зоною [17]. В ідеалі вони працюють віртуально.

Методи відображення інформації та взаємодії

Доповнена реальність (AR) дозволяє користувачеві зосередитися на інформації, що надається комп'ютером, та одночасно на реальному світі.

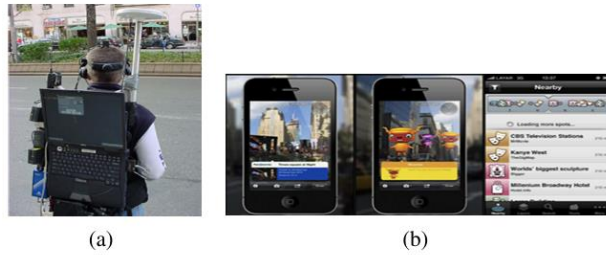


Рисунок 2.2 - Мобільна AR: (a) користувач із рюкзаком мобільної системи AR; (b) приклад AR-застосунку, що використовує мобільні -пристрої.

Інтерфейс користувача для візуальної мобільної доповненої реальності (AR) може поєднувати елементи, стабілізовані за екраном, стабілізовані за тілом та стабілізовані за світом. На рисунку 2.2 показано інтерфейс користувача з роботи авторів з мобільної доповненої реальності, сфотографований через оптичні прозорі дисплеї, що носяться на голові. Віртуальні прапорці та мітки – це об'єкти, стабілізовані за світом, що знаходяться у світовій системі координат і позначають точки інтересу в навколишньому середовищі. Вони відображаються у правильній перспективі для точки зору користувача, тому користувач може підходити до цих об'єктів та обходити їх так само, як і до фізичних об'єктів. Мітки спрямовані до користувача та зберігають свій розмір незалежно від відстані, щоб забезпечити читабельність. Синій та зелений рядки меню у верхній частині стабілізовані за розміром екрана, тобто вони займають однакове положення на екрані незалежно від того, куди дивиться користувач, як і конусоподібний вказівник у нижній частині екрана, який завжди вказує на поточний вибраний об'єкт світу. Два зображення на рисунку 2.3 (a) та (b) показують однакові параметри меню на місці, пов'язані з червоним прапорцем перед колонками, реалізовані двома різними способами. У частині (a), нашої початковій реалізації, меню було розташоване у вигляді кругового прапора зі стабілізацією за розміром екрана. Це спричиняло проблеми, коли користувач повертав голову під час вибору пункту меню. У дизайні, показаному в частині (b), меню є стабілізованим за розміром екрану елементом, пов'язаним з пов'язаним з ним прапорцем за допомогою Методи відображення інформації та

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						30
Змн.	Арк.	№ докум.	Підпис	Дата		

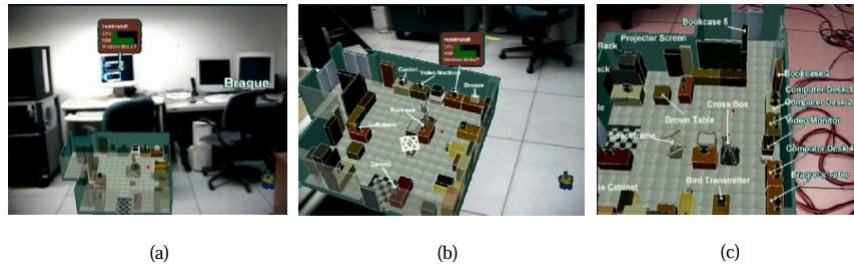


Рисунок 2.3 - Регулювання нахилу голови для WIM.

(а) Користувач дивиться прямо перед собою. (б) WIM збільшується, нахиляється та переміщується вгору по екрану, коли користувач нахиляє голову вниз. (в) Вигляд WIM майже зверху вниз, коли користувач дивиться ще нижче.

Доповнена реальність (AR) дозволяє користувачеві зосередитися на інформації, що надається комп'ютером, та одночасно на реальному світі. Інтерфейс користувача для візуальної мобільної доповненої реальності (AR) може поєднувати елементи, стабілізовані за екраном, стабілізовані за тілом та стабілізовані за світом. На рисунку 2.4 показано інтерфейс користувача з роботи авторів з мобільної доповненої реальності, сфотографований через оптичні прозорі дисплеї, що носяться на голові. Віртуальні прапорці та мітки – це об'єкти, стабілізовані за світом, що знаходяться у світовій системі координат і позначають точки інтересу в навколишньому середовищі. Вони відображаються у правильній перспективі для точки зору користувача, тому користувач може підходити до цих об'єктів та обходити їх так само, як і до фізичних об'єктів. Мітки спрямовані до користувача та зберігають свій розмір незалежно від відстані, щоб забезпечити читабельність. Синій та зелений рядки меню у верхній частині стабілізовані за розміром екрана, тобто вони займають однакове положення на екрані незалежно від того, куди дивиться користувач, як і конусоподібний вказівник у нижній частині екрана, який завжди вказує на поточний вибраний об'єкт світу. Два зображення на рисунку 2.2 (а) та (б) показують однакові параметри меню на місці, пов'язані з червоним прапорцем

перед колонками, реалізовані двома різними способами. У частині (а), нашій початковій реалізації, меню було розташоване у вигляді кругового прапора зі стабілізацією за розміром екрана. Це спричиняло проблеми, коли користувач повертав голову під час вибору пункту меню. У дизайні, показаному в частині (b), меню є стабілізованим за розміром екрану елементом, пов'язаним з пов'язаним з ним прапорцем за допомогою [18].

2.2 Методи комп'ютерного зору для мобільних AR-систем

Для середовищ, для яких геометричну модель неможливо легко отримати, або для аналізу, спрямованого на високодинамічні сцени, були запропоновані чисто візуальні підходи. Ці підходи зосереджені на аналізі локальних особливостей, таких як область інтересу навколо фіксованої області. Однак ці підходи потім вимагають фази навчання, на якій екземпляри позначаються людиною-анотатором.

У своєму підході SemantiCode використовують напівавтоматичний підхід, заснований на класифікації 2D-зображень (гістограма кольорів). Дані погляду аналізуються SemantiCode, а зразки оригінального відео з камери сцени, зосереджені навколо фіксації, представлені анотатору-людині. Анотатор позначає зразки, а SemantiCode крок за кроком створює базу даних прикладів, які використовуються для пропонування позначок у наступних презентаціях.

Представили Museum Guide 2.0, який використовує адаптований екстрактор ознак SIFT та підхід приблизного найближчого сусіда для класифікації зображень фіксованих об'єктів, отриманих з камери сцени, на основі попередньо визначеної бази даних потенційних цільових об'єктів. База даних має бути заповнена екземплярами зображень цільових об'єктів, знятих з кожної точки зору. Швидкість класифікації залежить від кількості ознак та розміру бази даних. Система могла працювати в режимі реального часу, коли обробка зображень була скорочена до менш ніж 25 кадрів на секунду на невеликому наборі об'єктів. Розширили цей підхід за допомогою просторово організованої

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						32
Змн.	Арк.	№ докум.	Підпис	Дата		

бази даних, щоб значно пришвидшити обробку. Запропонували подібний підхід до інтерактивного створення навчальних прикладів з їхнім попереднім етапом навчання *шляхом вивчення*.

Описані 2D-підходи можуть мати справу як зі стаціонарними, так і з рухомими об'єктами, що цікавлять. Однак вони вимагають інтенсивного етапу збору даних, і класифікації не гарантовано працюватимуть у режимі реального часу, оскільки вони залежать від розміру бази даних. Якість виявлення також сильно залежить від характеристик цільових об'єктів, і неможливо розрізнити кілька випадків.

Геометричні підходи

Геометричні підходи використовують 2D або 3D модель середовища для обчислення перетину лінії зору та об'єктів інтересу. Однак вони вимагають визначення положення та орієнтації системи відстеження погляду принаймні один раз для кожної фіксації (відстеження). Вони також вимагають, щоб геометрична модель була точно вирівняна з реальними цільовими об'єктами. Технічно, вирівнювання (часто також називається реєстрацією) та відстеження системи відстеження погляду можуть оброблятися однією системою.

Галузі взаємодії людини з комп'ютером, де повсюдно поширені точні 3D-геометричні моделі видимого середовища, – це віртуальна реальність та доповнена реальність [19]. Використовували погляд як метод взаємодії в захопливому середовищі віртуальної реальності, представленому через дисплей, встановлений на голові, для вибору об'єктів, вільно розташованих у 3D-просторі. Інші використовували погляд для співпраці у віртуальній реальності, для навчання виконанню завдань інспекції або для системних взаємодій. Усі ці підходи мають спільну рису: вони мають здатність класифікувати об'єкти інтересу в режимі реального часу, але для досягнення цієї мети потрібні штучні представлення цільових об'єктів.

Нещодавній підхід, який використовує подібні методи, був представлений. Вони змогли реконструювати так звані 3D-об'єми уваги над реальними об'єктами, але їм бракує геометричної моделі, і тому вони не можуть

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						33
Змн.	Арк.	№ докум.	Підпис	Дата		

класифікувати фіксації на цільових об'єктах, що цікавлять. Більш перспективний підхід був представлений: вони використовують установку, що складається з камери та Microsoft Kinect, для створення 3D-сканування цільового середовища. Потім вони можуть оцінити положення та орієнтацію даних мобільного відстеження очей, використовуючи створену модель як один великий 3D-маркер. Області в цій 3D-моделі потім можна анотувати та використовувати як області інтересу. Як додаткова перевага, 3D-модель також можна використовувати для візуалізації зібраних даних. Їхній підхід дуже підходить для таких приміщень, як супермаркети, де об'єкти інтересу мають розумний розмір і не щільно упаковані в просторі. Однак він вимагає інтенсивної фази підготовки, як вручну, так і обчислювально.

Десь посередині знаходиться робота, яка використовує погляд для взаємодії з системою доповненої реальності також у контексті технічного обслуговування. Вони пов'язують реальні та віртуальні моделі, але зосереджуються на взаємодії на основі погляду з віртуальними конструкціями.

Ще один цікавий підхід запропонували, який створюють 3D-модель середовища на льоту, використовуючи методи структури та руху з області самокалібрації та картографування [20]. Створені моделі є досить грубими, оскільки учасники зазвичай не рухаються так безперервно, як це потрібно для плавної генерації моделі. Але грубу структуру середовища все ще можна ідентифікувати, і кожен точку 3D-моделі можна пов'язати з екземплярами зображень сцени, які були зроблені під час фіксації уваги на вибраній області.

У своєму підході FixTag вони використовували трекер ознак на основі комп'ютерного зору для оцінки положення та орієнтації трекера очей у просторі. Їхня система вимагає ручного калібрування восьми опорних точок у вибраних ключових кадрах для встановлення опорної точки для відстеження. На основі відстежуваних ознак положення та орієнтація трекера очей потім інтерполюються на всі сусідні відеокадри. Області інтересу, які необхідно заздалегідь визначити у 3D відносно деяких опорних точок, потім відображаються на простір зображення відео з камери сцени системи

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						34
Змн.	Арк.	№ докум.	Підпис	Дата		

відстеження очей на основі обчисленого перетворення трека очей. Виявлені фіксації (2D у просторі зображення) потім класифікуються відповідно до відображених областей інтересу. Система FixTag працювала автономно менш ніж у режимі реального часу (2 секунди на кадр на MacBook Pro з частотою 2,5 ГГц). Класифікується лише один кадр на фіксацію.

Відстеження маркерів

Деякі з вищезгаданих підходів, а також запропонований нами підхід використовують маркери, видимі або невидимі (інфрачервоні), для інструментального дослідження середовища.

Інструментування середовища маркерами завжди викликає дискусію щодо того, чи змінює наявність таких маркерів поведінку учасників під час спостереження. Принаймні для інфрачервоних маркерів, що використовувалися Тобієм, не було виявлено жодного значного впливу на наявність маркерів.

Підхід EyeSee3D

У наших поточних дослідженнях ми прагнемо знайти недороге рішення для автоматичного 3D-аналізу погляду, яке можна було б застосовувати онлайн на будь-якій мобільній системі відстеження погляду з камерою сцени. Оскільки не багато систем пропонують комплект розробки програмного забезпечення для онлайн-доступу до даних відстеження погляду та камери сцени, наш підхід також підтримує офлайн-аналіз.

Відправною точкою нашого підходу є мобільна система відстеження погляду зі сценичною камерою (див. рис. 2.4). Вона виявляє зіниці людини за допомогою підключених очних камер. На основі цього можна розрахувати напрямок погляду та виявити 2D-фіксації. Вони представлені як координати на 2D-зображенні сценичної камери, тобто погляд людини відображається на пікселях на відеозображенні. Саме це зазвичай візуалізується у відео з поглядом та курсором. Крім того, 3D-вектор напрямку можна обчислити в системі координат, яка є відносною до положення та орієнтації, що також називається *поза* камери сцени. Наша поточна реалізація EyeSee3D базується на окулярах відстеження погляду SMI, проте загальний підхід застосовний до всіх інших

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						35
Змн.	Арк.	№ докум.	Підпис	Дата		

систем відстеження погляду на основі камери сцени.

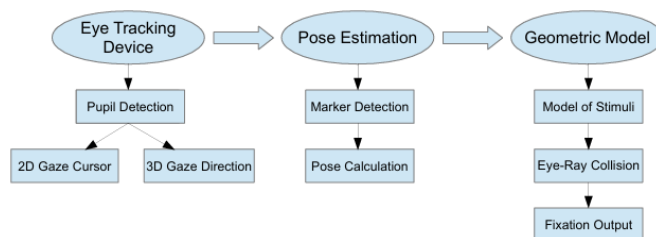


Рисунок 2.4 - Різні кроки підходу

Усі обчислення, пов'язані з поглядом, зазвичай виконуються пристроєм відстеження погляду. Зображення з камери сцени використовується для оцінки пози пристрою за допомогою відстеження маркерів [21]. Обидві інформації об'єднуються для ідентифікації об'єктів інтересу за допомогою raycasting у 3D-геометричній моделі.

Замість того, щоб дивитися на вміст відео з камери місця зйомки для ідентифікації об'єктів, що фіксуються, ідея, якої дотримується Eye-See3D, полягає у використанні 3D-векторного напрямку, що представляє лінію зору, та поєднанні його з оцінкою пози. Якщо камера місця зйомки відкалібрована, тобто її внутрішні параметри, такі як фокусна відстань, формат зображення та головна точка, визначені, можна перетворити відомі геометрії, виявлені на 2D-зображенні, на їх фактичну 3D-позицію. Іншими словами, можна розрахувати положення камери у світі відносно виявлених геометрій. Існує кілька таких геометрій, які можна легко виявити на зображеннях з камери. Ми обрали опорні маркери доповненої реальності, які мають високу контрастність та унікальний ідентифікаційний номер, закодований білими областями. Більше того, такі маркери з їхньою простою структурою можна виявляти в режимі реального часу.

Підсумовуючи, зв'язок камери сцени з її 3D-середовищем можна встановити шляхом виявлення опорних маркерів та обчислення трансформації пози відносно каліброваних параметрів камери. Поєднання цього з напрямком погляду призводить до 3D-променя погляду, який можна відправити в 3D-

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						36
Змн.	Арк.	№ докум.	Підпис	Дата		

середовище. Чого бракує для анотації фіксацій на 3D-об'єктах, що цікавлять, для аналізу погляду, так це їхнього положення та протяжності у світі. Для цього необхідно створити 3D-проксі-геометрію, що апроксимує об'єкти, що цікавлять, та прив'язати її відносно маркерів відстеження в геометричній моделі. Перевагою такої 3D-моделі є те, що її можна використовувати з будь-якої точки зору. Таким чином, аналіз погляду не прив'язаний до фіксованого розташування [22]. В EyeSee3D ми використовуємо стандарт W3C X3D для моделювання, який можна редагувати вручну, подібно до HTML. Він також підтримується багатьма інструментами графічного моделювання.

Віртуальну модель можна вирівняти зі світом, оскільки розташування маркерів відоме в обох. Потім її можна представити як накладання зображення сцени-камери: кожен об'єкт у віртуальному світі збігається за положенням та орієнтацією з відповідним реальним об'єктом. Якщо напрямок погляду тепер змодельовати як промінь, що починається з ока, об'єкти, які фіксуються, можна ідентифікувати, виявляючи зіткнення між променем та їх віртуальною проксі-геометрією. Потім система може повністю автоматично вивести появу фіксованого об'єкта розташовані відносно опорного маркера, їхнє положення неможливо відстежити за допомогою вбудованого алгоритму відстеження «зсередини назовні». Однак окремі маркери можна динамічно переміщувати, і кожна геометрія, пов'язана з таким маркером, буде відповідно оновлюватися. EyeSee3D можна розширити для підтримки додаткових алгоритмів відстеження для певних сценаріїв. Наведемо один приклад: у наступній налаштуванні ми додали відстеження обличчя для оновлення проксі-геометрії голови співрозмовника. Як альтернатива, рухи моделей також можна моделювати явно вручну, а потім використовувати офлайн-аналіз.

Наш метод також сумісний з іншими механізмами відстеження, такими як оптичне відстеження цілі зовні всередину або інерційне відстеження. Ці підходи можна додати для покращення оцінки пози пристрою відстеження очей і таким чином підвищити точність наступних кроків.

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						37
Змн.	Арк.	№ докум.	Підпис	Дата		

2.3 Інструментальні засоби та технічні вимоги до реалізації AR-додатків з використанням комп'ютерного зору

У цьому розділі ми детальніше розглянемо основні компоненти та інфраструктуру, необхідні для MARS. Ми розглянемо мобільні обчислювальні платформи, дисплеї для MARS, питання відстеження та реєстрації, моделювання навколишнього середовища, методи введення та взаємодії на носячих пристроях, бездротовий зв'язок та розподілене зберігання та доступ до даних. Ми коротко оглядаємо важливі історичні розробки в цих галузях та вказуємо на технології, які успішно використовувалися в прототипах MARS або мають великий потенціал для використання в майбутніх системах. Нарешті, ми підсумовуємо цей матеріал, описуючи гіпотетичний MARS найвищого класу, зібраний з найперспективніших компонентів, доступних на даний момент.

Мобільні обчислювальні платформи

Мобільні обчислювальні платформи за останні роки зазнали величезного прогресу в мініатюризації та продуктивності, і продаються за дедалі меншими цінами. Сьогодні високоякісні ноутбуки наздоганяють за обчислювальною потужністю доступні настільні рішення дуже швидко після появи на ринку нової моделі процесора. Тенденція до більшої мобільності чітко помітна. Ринок носимих технологій, хоча ще перебуває на початковому етапі, має зростаючу клієнтську базу в промисловості, уряді та військових. Носимі обчислювальні рішення для особистого використання тепер можна придбати з різних джерел.

Існує кілька факторів, що впливають на вибір обчислювальної платформи для мобільних досліджень у сфері доповненої реальності, включаючи необхідну обчислювальну потужність, форм-фактор та міцність всієї системи, споживання енергії, графічні та мультимедійні можливості, наявність портів розширення та інтерфейсу, доступну пам'ять та простір для зберігання даних, можливість оновлення компонентів, операційну систему та середовища розробки програмного забезпечення, наявність технічної підтримки та, нарешті, але не менш важливе, ціну. Цілком очевидно, що багато з них взаємозалежні. Чим

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						38
Змн.	Арк.	№ докум.	Підпис	Дата		

менший обчислювальний пристрій, тим менша ймовірність того, що він матиме найвищу обчислювальну потужність, графічні та мультимедійні можливості. Розширення та порти інтерфейсу зазвичай мають ціну збільшення розміру. Так само, як і можливість оновлення компонентів: якщо у вас є мініатюрна функціональність (наприклад, графіка) за допомогою спеціальних інтегральних схем, у вас більше немає можливості легко замінити цей компонент новою моделлю [23]. Крім того, потрібно багато зусиль та винахідливості, щоб зменшити будь-яку технологію до розміру, значно меншого, ніж пропонують конкуренти, тому можна очікувати, що таке обладнання продаватиметься за вищою ціною.

Проект MARS Колумбійського університету надає конкретний приклад деяких компромісів, пов'язаних зі складанням мобільної AR-платформи. Апаратна платформа для цього проекту, була зібрана з готових компонентів для максимальної продуктивності, можливості оновлення, простоти обслуговування та розробки програмного забезпечення. Цей вибір був зроблений за рахунок розміру та ваги прототипу системи, частини якої були встановлені на рамі рюкзака. З 1996 по 2002 рік кожен компонент прототипу був модернізований кілька разів до потужнішої або іншим чином більш вигідної версії, що було б неможливо, якби спочатку була обрана менша, більш інтегрована система. Обчислювальна потужність відповідала передовим технологіям мобільної обробки, починаючи від системи на базі Pentium 133 МГц у 1996 році і закінчуючи ноутбуком Pentium IV 2,2 ГГц у 2002 році. У той же час можливості 3D-графіки зросли з чіпа GLINT 500DTX із заявленою швидкістю заповнення близько 16,5 млн пікселів за секунду до NVIDIA Quadro4 500 Go із заявленими 880 млн пікселів за секунду.

Найменші платформи для носіння комп'ютерів, доступні наразі (Hybernaut Roma на базі Windows CE або OQO Ultra-Personal Computer з вищою продуктивністю, що незабаром стане доступним, Tiqit eightythree та Antelope Technologies Mobile Computer Core) забезпечують лише скромну графічну продуктивність, а їхні основні процесори не мають достатньої потужності для

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						39
Змн.	Арк.	№ докум.	Підпис	Дата		

програмного рендерингу складних 3D-сцен з інтерактивною швидкістю. Фактори, що впливають на вибір 3D-графічної платформи для мобільної доповненої реальності (AR), включають необхідну *графічну продуктивність*, *відео- та текстурну пам'ять*, підтримку *графічної бібліотеки* (OpenGL або Direct-X), наявність *стереодрайверів*, *енергоспоживання* та *ціну*. Найбільш практичним рішенням для мобільної AR-системи, яка може підтримувати складну інтерактивну 3D-графіку, є невеликі ноутбуки з інтегрованим 3D-графічним чіпом. Дисплей можна зняти, якщо комп'ютер використовується виключно з дисплеєм для близького огляду. Однак, з нашого досвіду роботи з Touring Machine, його можна добре використовувати в прототипних системах для налагодження та для забезпечення огляду для глядачів під час демонстрацій технологій.

Вимоги до певних застосувань можуть суттєво обмежити вибір обчислювальної платформи MARS. Наприклад, наприкінці 2002 року не було доступних інтегрованих рішень для носимих комп'ютерів, які б підтримували рендеринг та відображення складних графічних сцен у стерео форматі [24]. Системний розробник, який орієнтований на такі застосування, повинен або зібрати власне обладнання для створення рішення з малим форм-фактором, або вдаватися до найменшого доступного потужного ноутбука, який має достатню графічну продуктивність та графічний чіп, що підтримує стерео.

Мобільні системи не обов'язково мають функціонувати за схемою одного автономного пристрою, який носить із собою або носить на руках, генеруючи та представляючи всю інформацію користувачеві. Натомість можуть бути різні ступені «участі в середовищі», що використовують ресурси, які не обов'язково розташовані на тілі користувача. У найбільш орієнтованому на пристрій випадку вся інформація генерується та відображається на одному пристрої, який користувач носить або носить із собою [25]. Прикладами є портативні аудіоплеєри та кишенькові органайзери без можливості бездротового зв'язку.

Відходячи на один крок від підходу, орієнтованого на пристрій, функціональність може бути розподілена між кількома пристроями. Технології

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						40
Змн.	Арк.	№ докум.	Підпис	Дата		

бездротового зв'язку, такі як IEEE 802.11b або Bluetooth, стають у пригоді для обміну даними між різними пристроями. Наприклад, персональний органайзер або портативний комп'ютер можуть надсилати дані через бездротове з'єднання на аудіо/відеогарнітуру. Додавши мобільний телефон із підтримкою Bluetooth для глобального зв'язку, така комбінація становитиме повноцінне бездротове мобільне обчислювальне рішення. Не всі пристрої потрібно постійно носити з собою користувачем. Припустимо, що ми хочемо мінімізувати розмір та енергоспоживання портативного обчислювального обладнання. Не маючи обчислювальної потужності для генерування та обробки інформації, яка має відображатися, ми можемо перетворити портативний пристрій на так званий *тонкий клієнт*, покладаючись на зовнішні сервери для збору та обробки інформації та її передачі на портативний пристрій у вигляді потоку даних, який можна зручно представити з обмеженими доступними ресурсами.

У крайньому випадку, мобільному користувачеві не потрібно було б носити з собою *жодне* обладнання, і він все одно міг би користуватися мобільною доповненою реальністю (AR). Усі обчислення та сенсорні процеси могли б відбуватися в навколишньому середовищі. Сітку камер можна було б налаштувати таким чином, щоб кілька камер охоплювали будь-яке можливе місце, де може перебувати людина. Інформацію можна було б зберігати, збирати, обробляти та генерувати в мережі обчислювальних серверів, які не обов'язково повинні бути в полі зору або навіть поруч. Дисплеї, такі як гучномовці, відеостіни та проектоване відео, могли б надавати персоналізовану інформацію людям, які стоять або рухаються поруч. Однак інфраструктура, необхідна для такого сценарію, є досить складною. Завдання забезпечення роботи мобільної доповненої реальності в непередбачуваних середовищах вимагає рішень, ближчих до кінця спектра «один пристрій».

Дисплеї для мобільної доповненої реальності (AR)

Існують різні підходи до відображення інформації мобільній людині, і для цієї мети можна використовувати різноманітні типи дисплеїв: персональні портативні, наручні або головні дисплеї; екрани та спрямовані гучномовці,

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						41
Змн.	Арк.	№ докум.	Підпис	Дата		

вбудовані в навколишнє середовище; проекція зображення на довільні поверхні, і це лише деякі з них. Деякі з цих можливостей відображення також можна використовувати в комбінації.

Загалом, можна розрізняти дисплеї, які людина носить на тілі, та дисплеї, які використовують ресурси навколишнього середовища. Носимі аудіоплеєри та персональні цифрові органайзери використовують дисплеї, що належать до першої категорії, як і носимі комп'ютери з дисплеями, що носяться на голові. Прикладом другої категорії може бути персоналізована реклама, яка відображається на відеоекранах, повз які проходить людина. Для такого сценарію потрібна досить складна екологічна інфраструктура. Дисплеї повинні бути вбудовані в стіни та інші фізичні об'єкти, і вони повинні бути або оснащені датчиками, які можуть виявляти присутність конкретної особи, або вони можуть отримувати інформацію про відстеження перехожих через комп'ютерну мережу [26]. Таких середовищ поки що не існує поза межами дослідницьких лабораторій, але кілька дослідницьких груп почали досліджувати повсюдні середовища дисплеїв як частину систем «Розумного дому» або «Спільного телезачинування», таких як проект Microsoft Research EasyLiving або «Офіс майбутнього».

Ще один тип дисплеїв, який досліджується в дослідженнях доповненої реальності (AR), – це *проекційні дисплеї, що носяться на голові*. Цей тип дисплеїв, що носяться на голові, складається з пари мікродисплеїв, роздільників променя та мініатюрних проекційних лінз. Він вимагає стратегічного розміщення світловідбивного листового матеріалу в навколишньому середовищі. Дисплей, що носить на голові, проектує зображення у світ, і лише коли користувачі дивляться на ділянку світловідбивного матеріалу, вони бачать зображення, яке було надіслано з їхнього дисплея. Цей підхід спрямований на поєднання концепції фізичних поверхонь дисплея (у цьому випадку: ділянок світловідбивного матеріалу) з гнучкістю персоналізованих накладок за допомогою окулярів доповненої реальності. Унікальне персоналізоване зображення може бути згенеровано для кожної людини з групи людей, які

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						42
Змн.	Арк.	№ докум.	Підпис	Дата		

дивляться на один і той самий об'єкт зі світловідбивним покриттям, за умови, що їхні кути огляду не надто близькі один до одного.

Одним із перспективних підходів для мобільної доповненої реальності (AR) може бути поєднання різних технологій відображення. Дисплеї, що носяться на голові, забезпечують один із найбезпосередніших способів доступу до графічної інформації. Глядачеві не потрібно відводити погляд від об'єкта фокусування в реальному світі [27]. Безпосередність та конфіденційність персонального дисплея, що носить на голові, добре доповнюються високою читабельністю тексту портативних плазмових або РК-дисплеїв, а також можливостями спільної роботи дисплеїв настінного розміру. Наприклад, дослідження мобільної доповненої реальності в Колумбійському університеті на ранніх етапах експериментували з дисплеями, що носяться на голові, та портативними дисплеями, що використовувалися в синергетичній комбінації.

Як згадувалося в історичному огляді, концепція *прозорих дисплеїв комп'ютерної графіки, що носяться на голові*, сягає корінням у роботу Івана Сазерленда над 3D-дисплеєм, що носить на голові. Незадовго до цього, у 1957 році, Мортон Хейліг подав патент на дисплей, що носить на голові, оснащений двома кольоровими телевізійними блоками. У наступні роки було розроблено кілька дисплеїв, що носяться на голові, для досліджень у комп'ютерному моделюванні та військових цілях, включаючи роботу Тома Фернесса над системами відображення на голові для пілотів винищувачів.

Вибираючи дисплей для мобільної доповненої реальності (AR), який носить на голові, необхідно враховувати кілька факторів. Одним із найбільш суперечливих питань у дослідницькій спільноті доповненої реальності є вибір між оптичними прозорими дисплеями та прозорими відеодисплеями. Оптичні прозорі дисплеї прозорі, як і окуляри за рецептом або сонцезахисні окуляри. Вони використовують оптичні об'єднувачі, такі як дзеркальні розщеплювачі променя, для накладання згенерованого комп'ютером зображення поверх зображення, яке видно користувачеві.

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						43
Змн.	Арк.	№ докум.	Підпис	Дата		

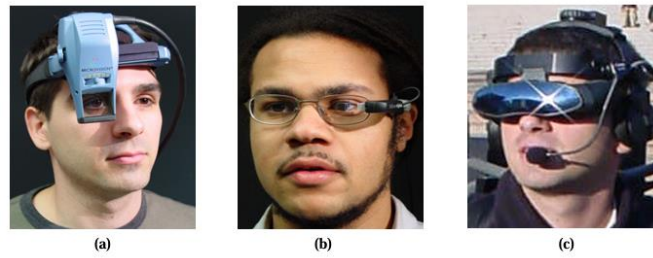


Рисунок 2.5 - Монокулярні та бінокулярні наголовні дисплеї з прозорою оптикою. (а) Microvision Nomad. (б) MicroOptical Clip-on. (в) Sony LDI-D100B (встановлений на кріплення, надане замовником).

На рисунку 2.5 (а) показано зображення, зняте через такі окуляри. На відміну від цього, прозорі відеодисплеї представляють більш непрямий, опосередкований вигляд навколишнього середовища. Одна або дві невеликі відеокамери, встановлені на дисплеї, який носиться на голові, захоплюють відеопотоки середовища перед користувачем, які відображаються на непрозорих екранах з відповідною оптикою прямо перед очима користувача. Комп'ютер може змінювати відеозображення перед тим, як воно буде надіслано на окуляри, для створення доповненої реальності (AR-накладок). Приклад показано на рисунку 2.5 (б). Недоповнені додатки для носимих комп'ютерів часто використовують монокулярні дисплеї. Навіть якщо дисплей непрозорий, користувач може бачити реальний світ незакритим оком [28]. Однак у такому випадку сприйняття справжнього поєднання комп'ютерного шару та реального світу може бути дещо складним.

Для мобільної роботи з доповненою реальністю автори цього розділу надають перевагу оптичним прозорим дисплеям. Ми вважаємо, що людина, яка пересувається в навколишньому середовищі, повинна мати можливість покладатися на свій повний природний зір. Хоча доповнена реальність може покращити зір, вона не повинна надмірно його погіршувати. На нашу думку, кілька недоліків сучасної технології прозорого відео перешкоджають її впровадженню в справді мобільних додатках: бачення реального світу з роздільною здатністю відео та тим самим малим кутом зору, що

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						44
Змн.	Арк.	№ докум.	Підпис	Дата		

використовується для графічних накладень, необхідність компенсації спотворень зображення, що вносяться камерами, ризик затримки відеопотоку на дисплей та проблеми безпеки, пов'язані з баченням світу виключно через камери.

За нашим досвідом, монокулярні дисплеї можуть давати прийнятні результати для AR, якщо дисплей прозорий, щоб користувачеві було легше поєднувати доповнене зображення з зображенням реального світу іншим оком, як у випадку з MicroVision Nomad. Більше поле зору також корисне. Однак важко обговорювати такі властивості дисплея окремо, оскільки на якість мобільних AR-презентацій впливає чимало факторів дисплея, серед яких монокулярний, біокулярний (для двох очей) та біокулярний (стерео), роздільна здатність, глибина кольору, яскравість, контрастність, поле зору, глибина фокусування, ступінь прозорості, вага, ергономіка та зовнішній вигляд. Споживання енергії є додатковим фактором при тривалому мобільному використанні.

Стереодисплеї можуть значно покращити враження від доповненої реальності (AR), оскільки віртуальні об'єкти можна краще сприймати на тій самій відстані, що й об'єкти реального світу, які вони анотують. Однак зауважте, що хоча стерео дозволяє відображати об'єкти з правильною невідповідністю для лівого/правого ока, всі доступні на даний момент дисплеї відображають графіку з однаковою видимою глибиною, і тому вимагають від очей глядача акомодатії на цій конкретній відстані, що призводить до конфлікту акомодатії та конвергенції. Наприклад, під час накладання віртуального об'єкта на реальний в оптичному прозорому дисплеї, якщо реальний об'єкт не розташований на цій конкретній фіксованій відстані, глядачеві потрібно налаштувати акомодатію, щоб побачити у фокусі або реальний, або віртуальний об'єкт.

Наразі можливості оптичних прозорих дисплеїв, що носяться на голові, досить обмежені. Якщо стереозвук є необхідністю, то можливості ще більш обмежені. Прототипи Колумбійського університету MARS протягом багатьох років використовували кілька оптичних прозорих дисплеїв зі стереозвуком, жоден з яких більше не представлений на ринку. На рисунку 2.5 (с) показано Sony LDI-D100B, дисплей, виробництво якого було знято з виробництва в червні

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						45
Змн.	Арк.	№ докум.	Підпис	Дата		

2000 року.

Дисплеї здебільшого й досі виглядають громіздкими та незграбними. Менші монокулярні дисплеї, такі як MicroOptical CO-1, зображений на рисунку 2.5 (b), або прототип Minolta «Forgettable Display», набагато менш помітні, але не забезпечують високих кутів поля зору, необхідних для справжнього занурення, а також яскравості, наприклад, Microvision Nomad. Тим часом виробники наполегливо працюють над удосконаленням та подальшою мініатюризацією оптики дисплеїв. Мікродисплеї сьогодні можна знайти в різноманітних продуктах, включаючи видошукачі для камер, дисплеї для мобільних телефонів та інших мобільних пристроїв, а також портативні відеопроєктори. Дисплеї для зйомки біля ока становлять зростаючий сегмент застосування на ринку мікродисплеїв [29]. Привабливість мобільної доповненої реальності (AR) залежить від подальшого прогресу в цій галузі.

Відстеження та реєстрація

Окрім технології відображення, найважливішим технологічним викликом для загальної мобільної доповненої реальності (AR) є відстеження та реєстрація. AR вимагає надзвичайно точного відстеження положення та орієнтації для вирівнювання або реєстрації віртуальної інформації з фізичними об'єктами, які потрібно анотувати. Важко переконати людей, що віртуальні об'єкти, згенеровані комп'ютером, насправді існують у тому ж фізичному просторі, що й об'єкти реального світу навколо нас. У контрольованих середовищах обмеженого розміру в закритих обчислювальних лабораторіях дослідникам вдалося створити середовища, в яких рух голови та рук людини можна відстежувати з достатньо високою просторовою точністю та роздільною здатністю, низькою затримкою та високою частотою оновлення, щоб створити досить реалістичні інтерактивні середовища комп'ютерної графіки, які, здається, співіснують з фізичним середовищем. Зробити те саме в загальному мобільному середовищі набагато складніше. У загальному випадку мобільних пристроїв не можна очікувати, що можна буде покладатися на будь-яку інфраструктуру відстеження в середовищі. Обладнання для відстеження має бути достатньо

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		46

легким для зносу, досить стійким до ударів та зловживання, а також функціональним у широкому спектрі умов навколишнього середовища, включаючи освітлення, температуру та погоду. За цих обставин наразі не існує ідеального рішення для відстеження, і ми не можемо очікувати його знаходження найближчим часом [30]. Доведеться йти на компроміси в продуктивності відстеження, а додатком доведеться адаптуватися.

Технологія відстеження постійно вдосконалювалася з часів появи комп'ютерної графіки з відстеженням головних рухів. Оригінальний дисплей Сазерленда, що носився на голові, відстежувався механічно за допомогою стельового обладнання, і через все обладнання, підвішене до стелі, його жартома називали «Дамокловим мечем». Сазерленд також досліджував використання ультразвукового трекера для головних рухів. Впровадження магнітного трекера наприкінці 1970-х років мало великий вплив на дослідження віртуальної та доповненої реальності (VR) і доповненої реальності (AR), і та сама технологія, у вдосконаленому вигляді, використовується й сьогодні. Протягом 1990-х років стали доступними комерційні гібридні системи відстеження, засновані на різних технологіях, які досліджувалися окремо в експериментальних системах відстеження протягом попередніх десятиліть, таких як ультразвукове, магнітне та оптичне відстеження положення, а також інерційне та магнітометричне відстеження орієнтації. Що стосується систем глобального позиціонування, ідея сучасної системи GPS NAVSTAR народилася в 1973 році, перший робочий супутник GPS було запущено в 1978 році, а створення сузір'я з 24 супутників було завершено в 1993 році. Супутники для російського аналога сузір'я *Глонасс* запускалися з 1982 року. Європейський Союз планує запустити окрему систему GPS з 30 супутників під назвою *Galileo*.

У решті цього підрозділу ми розглянемо технології відстеження, які найбільше підходять для мобільної доповненої реальності (AR). Для більш повного огляду технологій відстеження для AR та VR ми відсилаємо читача до існуючих оглядів технологій та методів відстеження руху, або нещодавнього спеціального випуску журналу, присвяченого відстеженню.

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						47
Змн.	Арк.	№ докум.	Підпис	Дата		

Візуальну реєстрацію віртуальних та фізичних об'єктів можна досягти кількома способами. Один поширений підхід полягає у визначенні пози голови людини в деякій глобальній системі координат та співвіднесенні її з комп'ютерною моделлю поточного середовища. Зауважте, що в цьому випадку комп'ютерну модель середовища необхідно створити на етапі, який називається *моделюванням середовища*. Ця модель повинна використовувати ту саму глобальну систему координат, що й система відстеження, або ж необхідно знати необхідне перетворення. Визначення положення та орієнтації об'єкта часто називають відстеженням із шістьма ступенями свободи (6DOF) для шести параметрів, що вимірюються: положення по осях x, y та z, а також орієнтація за кутами рихання, тангажу та крену.

Абсолютне положення та орієнтація голови користувача та фізичних об'єктів, що потребують анотації, не обов'язково мають бути відомі. В одному з найпряміших підходів до візуальної реєстрації камери спостерігають за певними унікальними орієнтирами (наприклад, штучними маркерами) в навколишньому середовищі. Якщо параметри перегляду камери (положення, орієнтація, поле зору) збігаються з параметрами перегляду дисплея (наприклад, тому що дисплей відображає зображення з камери, як у випадку прозорих відеодисплеїв), і стереографіка не використовується, віртуальні анотації можна вставляти безпосередньо в піксельних координатах без необхідності встановлення точного геометричного співвідношення між маркером і камерою. З іншого боку, якщо точне розташування орієнтирів у навколишньому середовищі відомо, для оцінки пози камери можна використовувати методи комп'ютерного зору. Використання камер, встановлених на дисплеї, разом із розпізнаванням орієнтирів іноді називають *відстеженням із замкнутим циклом*, при якому точність відстеження може бути скоригована до найближчого пікселя, якщо зображення з камери та графічне відображення збігаються. Це відрізняється від *відстеження з відкритим циклом*, яке намагається узгодити віртуальні анотації з фізичними об'єктами в реальному світі, спираючись виключно на сприйняту позу людини в 6 ступенях свободи та комп'ютерну модель середовища. Будь-які неточності в

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						48
Змн.	Арк.	№ докум.	Підпис	Дата		

пристроях відстеження або геометричній моделі призведуть до незначного зміщення анотації від її передбачуваного положення відносно фізичного світу [31].

Важливим критерієм для мобільного AR-відстеження є те, скільки обладнання для відстеження потрібно на тілі користувача та в навколишньому середовищі. Очевидна мета - носити якомога менше обладнання та не вимагати жодної підготовки середовища. Зауважте, що така система, як GPS, відповідає цій вимозі для всіх цілей, навіть якщо потрібне «підготовлене середовище» в глобальному масштабі у вигляді супутникового сузір'я. Кілька підходів до відстеження вимагають певних знань про середовище. Щоб створити будь-які корисні AR-анотації, об'єкти, які потрібно анотувати, повинні бути змодельовані або *геоприв'язані* в абсолютному вираженні, або їхнє розташування має бути можливим визначити за відомим зв'язком із попередньо вибраними та ідентифікованими орієнтирами.

Точність відстеження, необхідна для мобільної доповненої реальності (AR), дуже залежить від застосування та відстані до об'єктів, які потрібно анотувати. Якщо ми анотуємо приблизні контури будівель, ми можемо дозволити собі певну похибку реєстрації. Намагаючись точно визначити місцезнаходження певного вікна, ми повинні бути точнішими. Коли помилки реєстрації вимірюються як відстань на екрані між проектованою фізичною цільовою точкою та точкою, де малюється анотація, діє таке спостереження: чим далі об'єкт, який потрібно анотувати, тим менше помилок відстеження положення впливають на точність реєстрації. Зворотне вірно для помилок відстеження орієнтації. Оскільки більшість цілей у зовнішній мобільній доповненій реальності, як правило, знаходяться на певній відстані від глядача, можна припустити, що помилки відстеження орієнтації набагато більше сприяють загальній неточності реєстрації, ніж помилки відстеження положення. Оскільки немає автономних датчиків, які б забезпечували загальне надійне відстеження 6 ступенів свободи в непідготовлених зовнішніх умовах, мобільні системи доповненої реальності зазвичай вдаються до гібридних підходів, часто

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						49
Змн.	Арк.	№ докум.	Підпис	Дата		

використовуючи окремі механізми для відстеження положення та орієнтації.



Рисунок 2.6 – Приклад комп'ютерного зору для AR-додатків

Відстеження положення за допомогою GPS є природним кандидатом для використання на відкритому повітрі, оскільки воно функціональне в глобальному масштабі, якщо можна приймати сигнали щонайменше з чотирьох супутників [32]. Хоча використання GPS-навігації довгий час обмежувалося районами, що забезпечують пряму видимість супутників, так звана *допоміжна GPS* (A-GPS) у багатьох випадках обходить це обмеження. A-GPS використовує всесвітню опорну мережу серверів і базових станцій для наземної трансляції сигналу. У поєднанні з великою кількістю паралельних кореляційних схем прийому в мобільному GPS-приймачі, область відстеження може бути розширена на багато раніше не виявлених районів, таких як міські каньйони та приміщення, де сигнал достатньо сильний.

Звичайна GPS без вибіркової доступності має точність приблизно 10-15 метрів. GPS, що використовує систему доповнення широкого діапазону (WAAS), зазвичай має точність 3-4 метри в США та інших країнах, які впроваджують цю технологію. Диференціальна GPS зазвичай дає оцінку місцезнаходження з точністю приблизно 1-3 метри з локальною базовою станцією. Кінематична GPS у реальному часі (RTK GPS) з роздільною здатністю неоднозначності фази несучої може давати оцінки місцезнаходження з точністю до сантиметра. Останні два варіанти вимагають наявності поблизу базової станції, з якої диференціальний сигнал корекції помилок може бути надсиланий на

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						50
Змн.	Арк.	№ докум.	Підпис	Дата		

роумінговий пристрій. Тому в такому випадку вже не можна говорити про невідповідне середовище. Однак доступні комерційні диференціальні послуги, базові станції яких охоплюють більшу частину Північної Америки. Протягом тривалого часу комерційні диференціальні GPS-приймачі забезпечували частоту оновлення до 5 Гц, що є неоптимальним для відстеження швидкого руху людей або об'єктів. Новіші продукти забезпечують частоту оновлення до 20 Гц. Ще одна система відстеження положення, що застосовується для мобільної доповненої реальності (AR) на великій відстані, передбачає обчислення місцезнаходження людини на основі показників якості сигналу бездротової мережі IEEE 802.11b (WiFi). Це, очевидно, також вимагає розгортання обладнання в середовищі, в цьому випадку точок доступу WiFi, але якщо така бездротова мережа є обраною комунікаційною технологією для мобільної системи AR, система позиціонування може слугувати додатковою перевагою. Кілька дослідницьких проектів і принаймні один комерційний продукт досліджують цю концепцію. Система RADAR використовує для цієї мети мультилатерацію та попередньо обчислені карти сили сигналу, тоді як Використовують підхід байєсівських мереж. Досяжна роздільна здатність залежить від щільності точок доступу, розгорнутих для формування бездротової мережі. Екахау пропонує програмний продукт, який дозволяє відстежувати положення пристроїв з підтримкою WiFi після ручного збору/калібрування даних.

У MARS часто використовуються два додаткові засоби визначення положення, здебільшого як частина гібридних систем відстеження: інерціальні датчики та підходи на основі зору. Акселерометри та гіроскопи – це автономні або *безджерельні* інерціальні датчики. Їхня основна проблема – дрейф. Вихідні дані акселерометрів необхідно інтегрувати один раз за часом, щоб відновити швидкість, і двічі, щоб відновити положення. Отже, будь-яке погіршення продуктивності необроблених даних призводить до швидкого зростання помилок у отриманій оцінці положення. На практиці цей підхід до оцінки положення може використовуватися лише протягом дуже коротких періодів часу

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						51
Змн.	Арк.	№ докум.	Підпис	Дата		

між оновленнями, зібраними з більш надійного джерела. Інерціальні датчики також можуть використовуватися для виявлення кроку пішохода. Це принцип функціонування *крокомірів*, які в поєднанні з точною інформацією про напрямок руху можуть забезпечити практичний метод *відліку місця*.

Підходи на основі зору є перспективним варіантом для оцінки пози в 6 ступенях свободи в загальному мобільному середовищі. На окулярах встановлена одна або дві крихітні камери, щоб комп'ютер міг приблизно бачити те, що бачить користувач. Методи зору на основі моделі вимагають точної моделі середовища з відомими орієнтирами, які можна розпізнати на зображеннях. На противагу цьому, алгоритми зіставлення рухів відстежують динамічно вибрані ключові точки вздовж послідовності зображень, що призводить до відносних, а не абсолютних рішень відстеження, що означає, що необхідно встановити подальшу реєстрацію системи координат послідовності зображень з фізичним світом, щоб забезпечити 3D-графічне накладання. Одночасна реконструкція руху камери та геометрії сцени можлива, але такі обчислення є дуже обчислювально дорогими, а існуючі алгоритми вимагають «пакетного коригування пакета», глобального офлайн-обчислення по всій послідовності зображень. Нарешті, методи відстеження ознак на основі 2D-зображень вимірюють так званий «оптичний потік» між наступними відеозображеннями. Такі методи порівняно швидкі, але самі по собі не можуть оцінити 3D-рух камери. Можливі комбінації всіх цих підходів. Нещодавні дослідження дають багатообіцяючі результати для деяких тестових сценаріїв. Однак, загалом, алгоритмам комп'ютерного зору все ще бракує надійності та вони вимагають настільки великого обсягу обчислень, що суто візуальні рішення для загального відстеження в режимі реального часу все ще недосяжні. Наразі найбільший потенціал демонструють гібриди відстеження на основі зору та інших сенсорних технологій.

Відстеження орієнтації також значно виграє від гібридних підходів. Основними технологіями, доступними для визначення орієнтації, є електромагнітні компаси (магнітометри), гравітаційні датчики нахилу

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						52
Змн.	Арк.	№ докум.	Підпис	Дата		

(інклінометри) та гіроскопи (механічні та оптичні). Були розроблені гібридні рішення як у вигляді комерційних продуктів, так і дослідницьких прототипів. Датчики орієнтації IS300 та InertiaCube2 від InterSense поєднують три мікроелектромеханічні гіроскопи, три акселерометри (для прогнозування руху) та електромагнітний компас в одному невеликому інтегрованому датчику. Представили гібридний трекер, який поєднує ретельно калібрований компас і датчик нахилу з трьома гіроскопами швидкості. Розширили цю систему алгоритмом зору, що відповідає руху, який, однак, не працював у режимі реального часу. Представив метод корекції на основі зору, що базується на порівнянні силуету лінії горизонту з моделлю місцевої географії. Застосували метод зіставлення шаблонів на вручну вибраних орієнтирах в алгоритмі реального часу, який коригує дрейф орієнтації високоточного волоконно-оптичного гіроскопа.

Підсумовуючи, проблема відстеження пози людини для загальних цілей мобільної доповненої реальності (AR) є складною проблемою, для якої немає єдиного найкращого рішення. Гібридні підходи до відстеження наразі є найперспективнішим способом вирішення труднощів, що виникають у загальних середовищах мобільної доповненої реальності в приміщенні та на вулиці.

2.4 Висновок по розділу

У другому розділі було досліджено методи та інструменти комп'ютерного зору в доповненій реальності. Проведено порівняльний аналіз технологій доповненої та віртуальної реальності, що дозволило визначити особливості їх застосування у мобільних системах. Розглянуто методи комп'ютерного зору для AR-додатків, зокрема алгоритми відстеження, розпізнавання об'єктів і аналізу зображень. Також досліджено інструментальні засоби та технічні вимоги до реалізації AR-додатків із використанням комп'ютерного зору. У результаті визначено ефективні технології та підходи для створення сучасних мобільних AR-систем.

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						53
Змн.	Арк.	№ докум.	Підпис	Дата		

РОЗДІЛ 3. РОЗРОБКА ТА ІНТЕГРАЦІЯ КОМП'ЮТЕРНОГО ЗОРУ В МОБІЛЬНИЙ AR-ДОДАТОК

3.1 Проектування архітектури мобільного AR-додатку з використанням комп'ютерного зору

Під час аналізу та проектування системи необхідно створити її повні та узгоджені моделі. Модель – це набір взаємопов'язаних абстрактних елементів з можливим зазначенням їх властивостей, поведінки та зв'язків між ними. Для побудови діаграм було використано Edraw Max [39].

Діаграма варіантів використання. Метою розробки діаграми варіантів використання є визначення загальних меж системи та предметної області; формування загальних вимог до функціональної поведінки системи; розробка початкової моделі системи та подальша деталізація у вигляді логічної та фізичної моделей [17]. Діаграма варіантів використання представлена на рисунку 3.1.

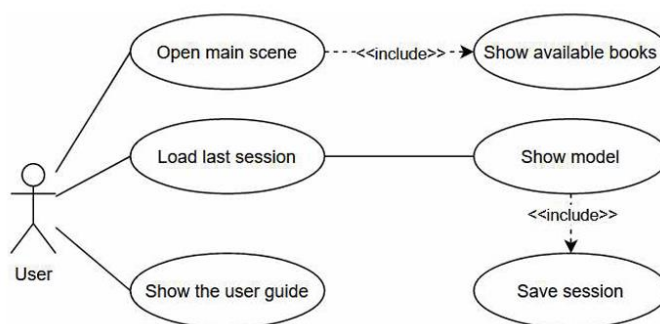


Рисунок 3.1 - Діаграма варіантів використання мобільної системи
комп'ютерного зору для освіти з використанням технології доповненої
реальності (AR).

Користувач має можливість отримати інформацію про модель; завантажити останній сеанс, що означає завантажити інформацію про останню переглянута модель; зберегти поточний сеанс; переглянути посібник користувача [33]. Отримання інформації про модель передбачає вибір книги з

поточної бази даних книг, для яких доступні моделі.

Діаграма класів. Діаграма класів необхідна для представлення статичної структури моделі системи. Діаграма класів показує зв'язки між сутностями системи, їхню внутрішню структуру та типи зв'язків. Діаграма класів, представлена на рисунку 3.2, відображає структуру класів системи.

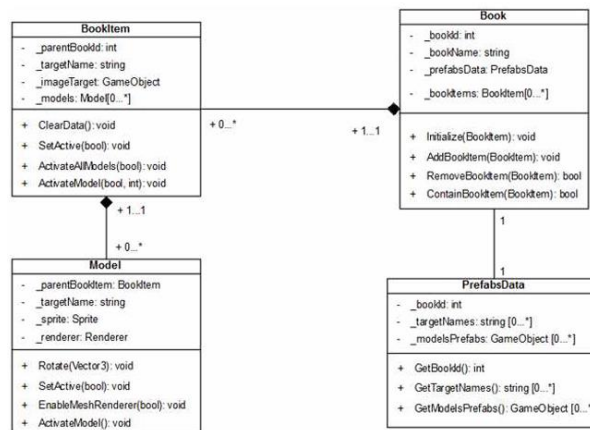


Рисунок 3.2 - Діаграма класів мобільної системи комп'ютерного зору для навчання з використанням технології доповненої реальності (AR).

Діаграма класів містить клас «Book», який містить посилання на інші об'єкти. Метод «Initialize()» цього класу генерує масив об'єктів типу «BookItem». Клас «BookItem» використовується для зберігання посилань на об'єкти, які безпосередньо є моделями, що відображаються користувачам. Його метод «ClearData()» дозволяє скидати всі поля класу. Клас має можливість активувати та деактивувати всі моделі цього об'єкта. На діаграмі показано зв'язки між основними класами системи та їх множинність.

Функціональна модель описує функції системи, можливі варіанти її використання; може містити інформацію про циркуляцію інформації в системі, об'єкти та сутності, що взаємодіють із системою; може бути динамічною або статичною моделлю. Побудова моделі починається з розробки контекстної IDEF0-діаграми (рисунок 3.3).

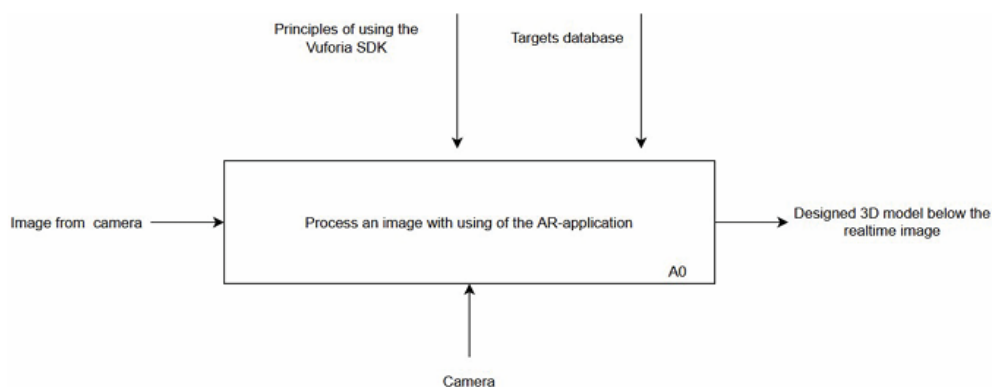


Рисунок 3.3 - Діаграма контексту верхнього рівня А-0 мобільної системи комп'ютерного зору для освіти з використанням технології AR.

Під час відкриття програми користувач повинен вибрати потрібний пункт меню, що відповідає книзі, яку він читатиме. Далі відбувається процес сканування відеопотоку та ідентифікації цілей. Якщо ціль знайдена в базі даних, виконується пошук інформації про цей клас та її відображення на екрані пристрою.

Моделювання навколишнього середовища

Для цілей доповненої реальності часто корисно мати доступ до геометричних моделей об'єктів у фізичному середовищі. Як згадувалося вище, одне з застосувань таких моделей - реєстрація. Якщо ви хочете анотувати вікно в будівлі, комп'ютер повинен знати, де це вікно розташоване відносно поточного положення та поля зору користувача. Наявність детальної ієрархічної 3D-моделі будівлі, включаючи такі елементи, як підлоги, кімнати, двері та вікна, надає комп'ютеру гнучкість у відповідях на такі запитання. Деякі методи відстеження, такі як підходи комп'ютерного зору на основі моделей, явно спираються на ознаки, представлені в більш-менш детальних моделях середовища відстеження [30]. Геометричні комп'ютерні моделі також використовуються для визначення *перекриття* відносно поточного ракурсу спостерігача. Наприклад, якщо частини будівлі перед спостерігачем перекриваються іншими об'єктами, лише незакриті частини будівлі повинні бути анотовані назвою будівлі, щоб уникнути плутанини для спостерігача щодо того, який об'єкт анотований.

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						56
Змн.	Арк.	№ докум.	Підпис	Дата		

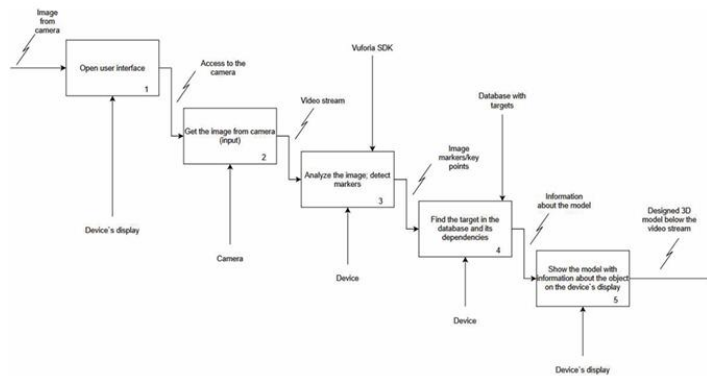


Рисунок 3.4 - Діаграма декомпозиції A-0 мобільної системи комп'ютерного зору для освіти з використанням технології доповненої реальності (AR).

Для цілей, згаданих вище, модель середовища не обов'язково має бути фотореалістичною. Можна не враховувати матеріали, текстури та, можливо, навіть геометричні деталі. Фактично, при відстеженні на основі моделі часто використовується лише «хмара» непов'язаних 3D-точок вибірки. Більш реалістичні геометричні моделі реальних структур, часто використовуються для анотацій або для огляду реального середовища. Наприклад, будівля, яка прихована від поля зору користувача, може бути відображена в її точному прихованому місці за допомогою доповненої реальності (AR), що дозволяє користувачеві, в певному сенсі, бачити крізь стіни.

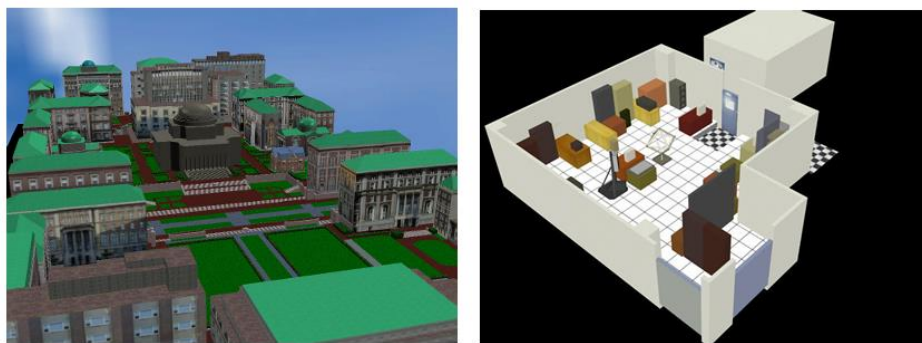


Рисунок 3.5 - Моделювання середовища: (а) Модель університетського містечка. (б) Модель лабораторії.

Людині, яка шукає певну будівлю, може бути показана її віртуальна

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						57
Змн.	Арк.	№ докум.	Підпис	Дата		

версія на екрані доповненої реальності (AR). Отримавши уявлення про форму та текстуру будівлі, людина може спостерігати, як модель рухається в правильному напрямку, доки вона не співпаде з реальною будівлею у фізичному просторі. Користувачеві може бути представлена тривимірна карта навколишнього середовища, щоб отримати огляд навколишнього середовища з висоти пташиного польоту. На рисунках показано приклади таких *світів у мініатюрі* (WIM), що використовуються в доповненій реальності (AR).

Створення 3D-моделей великих середовищ є саме по собі дослідницьким викликом. Можуть бути застосовані автоматичні, напівавтоматичні та ручні методи, серед яких 3D-реконструкція з супутникових знімків та аерофотознімків, 3D-зображення за допомогою лазерних далекомірів, реконструкція з набору частково перекриваючих фотографій, геодезична зйомка за допомогою тахеометрів та інших телеметричних інструментів, а також ручна реконструкція за допомогою програмного забезпечення для 3D-моделювання. Навіть сама доповнена реальність (AR) може бути використана для цілей моделювання. Огляд 3D-реконструкції синтетичного середовища. Моделі на рисунку 3.5 були реконструйовані шляхом екструзії контурів 2D-карти кампусу Колумбійського університету та нашої дослідницької лабораторії, ручного вдосконалення отриманих моделей та вибіркового текстурування їх за допомогою фотографій, зроблених з різних стратегічних позицій. WIM показує 3D-модель вибраної будівлі (Dodge Hall) у контексті аерофотознімка поточного середовища [31].

3D просторові моделі можуть бути як завгодно складними. Розглянемо, наприклад, завдання повного моделювання великої міської території, аж до рівня водопровідних труб та електричних кіл у стінах будівель. Існують значні дослідницькі проблеми, пов'язані з моделюванням, а також з організацією та зберіганням таких даних у просторових базах даних та структурах даних, оптимізованих для конкретних запитів. Зрештою, моделювання навколишнього середовища не закінчується статичною моделлю геометрії. Більшість середовищ є динамічними: зміни в геометричних моделях (через рухомі об'єкти, будівництво або руйнування) необхідно відстежувати та відображати в моделі

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						58
Змн.	Арк.	№ докум.	Підпис	Дата		

навколишнього середовища. Бази даних можуть потребувати досить швидкого змінювання, залежно від рівня деталізації.

3.2 Програмна реалізація системи комп'ютерного зору в AR-додатку

Наразі існує кілька фреймворків для роботи з AR, таких як Wikitude, EasyAR, ARToolkit, Vuforia. Ми проаналізуємо їх, щоб визначити оптимальний фреймворк для розробки мобільних додатків для освіти з використанням технології AR.

Оскільки застосунок буде базуватися на розпізнаванні цілей та 3D-моделюванні, необхідно, щоб фреймворки дозволяли працювати з різними налаштуваннями для цілі. Wikitude [40] дозволяє накладати лише статичні об'єкти доповненої реальності. EasyAR не надає жодних можливостей для легкого та швидкого використання цього фреймворку. Розробник має лише SDK та невелику документацію, яка описує основні принципи розпізнавання об'єктів. Документація охоплює лише основні принципи та специфікацію класу. ARToolkit – це бібліотека з відкритим кодом, але вона давно не отримувала оновлень, і всі приклади, що надаються з бібліотекою, неможливо запустити. Vuforia має власну онлайн-студію, яка генерує базу даних зображень цілей та показує точки, в яких відбувається розпізнавання. На відміну від Wikitude, Vuforia надає можливість завантажувати зображення, прямокутний паралелепіпед та циліндр як цілі. Google також має власну бібліотеку доповненої реальності ARCore, але вона зараз попередньо доступна та може бути нестабільною. Після аналізу всієї інформації та врахування доступного обладнання для тестування програми було вирішено обрати бібліотеку Vuforia.

Щоб мати змогу користуватися бібліотекою Vuforia, вам потрібно зареєструватися на офіційному вебсайті Vuforia. Після цього ви можете отримати ліцензію на проект [11].

Оскільки в проекті будуть використовуватися попередньо визначені зображення, необхідно створити базу даних з ними на стороні Vuforia. Можна

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						59
Змн.	Арк.	№ докум.	Підпис	Дата		

додавати різні типи маркерів. У нашому випадку буде використовуватися лише один тип маркера, яким є малюнок. Під час завантаження зображень до бази даних необхідно вказати ширину зображення, а також назву маркера. Після завантаження зображення визначаються його ознаки, за якими маркеру надається певний рівень (рейтинг). Чим вищий бал, тим більша ймовірність того, що маркер буде розпізнано. Приклад виділення ознак на зображенні показано на рисунку 3.6.

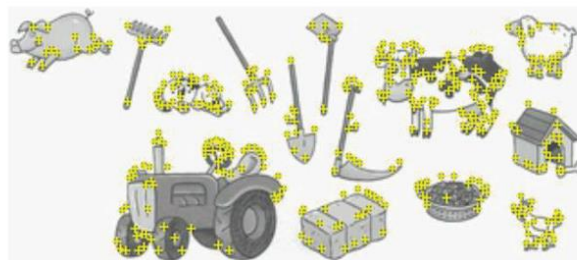


Рисунок 3.6 - Приклади об'єктів із виділеними елементами для використання в доповненої реальності.

Створена база даних завантажується на комп'ютер та додається до проекту. У сцені налаштовується взаємодія віртуальних об'єктів та цільових зображень. Найпростіший спосіб – прикріпити 3D-модель до об'єкта ImageTarget. Коли доповнена камера знаходить ціль, з'являється об'єкт ImageTarget із прикріпленим до нього 3D-об'єктом. Приклад представлено на рисунку 3.7.

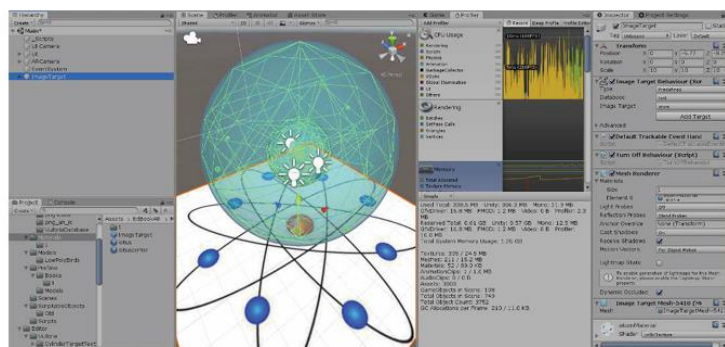


Рисунок 3.7 - Вікно редактора Unity із ImageTarget та 3D-об'єктом на сцені

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		60

Прості об'єкти можна створювати за допомогою алгоритмів для побудови сітки об'єкта. Кожна сітка складається з вершин і трикутників, які вони утворюють (рисунок 3.8 (а)). на рисунку 3.8 (b) показано куб як AR-об'єкт.

Складні об'єкти були попередньо завантажені в проєкт. Наприклад, сітка оленя показана на рисунку 3.8.

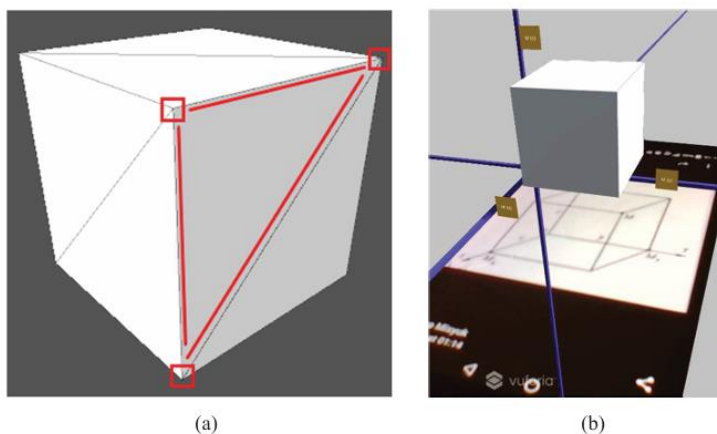


Рисунок 3.8 - Побудова 3D-моделей: куб (режим каркаса) із виділеними вершинами (а) та куб як об'єкт доповненої реальності (б)

Інтерфейс користувача мобільного програмного застосунку для освіти з використанням технології AR містить головне меню, меню вибору книги та меню вибору моделі, якщо для етикетки є кілька попередньо створених моделей [32]. Меню вибору моделі має список кнопок у правій частині екрана (Рисунок 3.9).

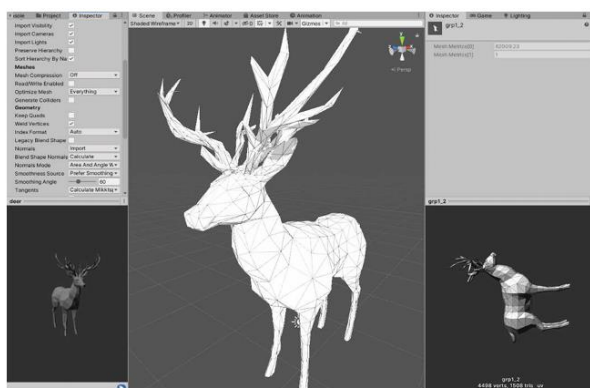


Рисунок 3.9 - Складні об'єкти

Користувач мав можливість переглядати інформацію про модель різними мовами. Крім того, користувач міг прослухати записану аудіо доріжку для цієї моделі. Приклад представлено на рисунку 3.9 (у нашому випадку як варіант українською мовою про віруси). Приклади застосування мобільної системи комп'ютерного зору для навчання з використанням технології доповненої реальності (AR) показано на рисунку 3.10.

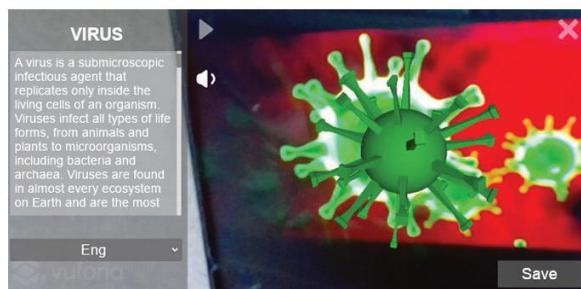


Рисунок 3.10 - Модуль Vuforia розпізнав об'єкт.

Проект має великий потенціал для розвитку та використання в освіті. Додаток можна використовувати для навчання дітей дошкільного віку. У цьому випадку книги можуть мати багато зображень тварин, рослин, об'єктів навколишнього середовища. Додаток може навчити дітей взаємодіяти з такими об'єктами. У сучасних школах можна використовувати цю систему на уроках біології, хімії, фізики, геометрії тощо. Під час сканування малюнка з дошки учень може отримати 3D-модель об'єкта, що вивчається. За наявності анімації для цієї моделі швидкість вивчення поняття/об'єкта учнем може зрости, оскільки він побачить принцип роботи цього об'єкта на дисплеї. Цей розроблений додаток доповненої реальності також можна використовувати у кіноіндустрії, робототехніці, освіті, медицині, транспорті тощо. Використовуючи іншу базу даних та заздалегідь завантаживши її в додаток, ви можете використовувати додаток під час навчання студентів-медиків. Наприклад, під час сканування зображення людського тіла органи та їхні назви можуть з'являтися у відповідних місцях мовою користувача та латиною; цей додаток можна використовувати для ідентифікації конкретного захворювання, яке візуально проявляється на тілі

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						62
Змн.	Арк.	№ докум.	Підпис	Дата		

пацієнта.

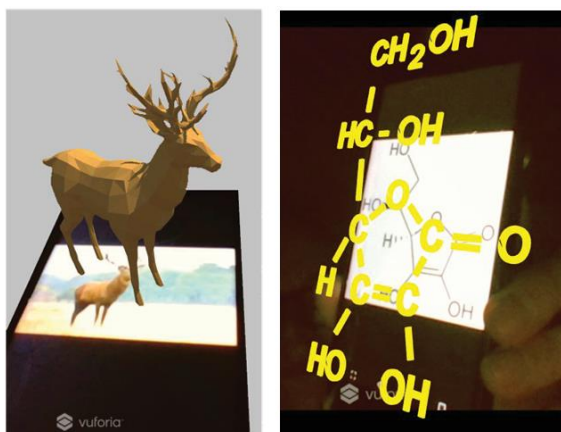


Рисунок 3.11 - Приклади застосування розробленої мобільної системи комп'ютерного зору

3.3 Тестування та оцінка ефективності алгоритмів відстеження та розпізнавання

Підхід EyeSee3D, представлений вище, був розроблений та оцінений у контексті експерименту, що аналізує поведінку погляду під час особистої взаємодії двох співрозмовників.

В експерименті співрозмовники сидять один навпроти одного на кінцях столу (див. рис. 3.12) . На столі п'ятьма рядами розташовано 23 фігурки LEGO Duplo™, кожна з яких звернена до одного з учасників. Фігурки мають певні риси, видимі лише для одного з учасників. Один з учасників одягає мобільні окуляри для відстеження погляду, оснащені камерою, яка записує поле зору. Кожна спроба починається з усного опису шуканої фігурки. Завдання для обох учасників полягає в тому, щоб знайти фігурку, а той, хто шукає, повинен домовитися зі співрозмовником, яка це фігурка. Дослідження проводилося з 12 парами учасників, і кожна пара мала ідентифікувати 20 фігурок.

Положення фігур є статичним, що дозволяє нам геометрично моделювати сценарій без урахування динамічно рухомих об'єктів. Фігури в установці є 3D-об'єктами, які розташовані по глибині та можуть частково перекривати один

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						63
Змн.	Арк.	№ докум.	Підпис	Дата		

одного залежно від перспективи глядача. Через ці перекриття підходи, обмежені 2D-площинами, не здатні надійно автоматично анотувати фіксації [34].

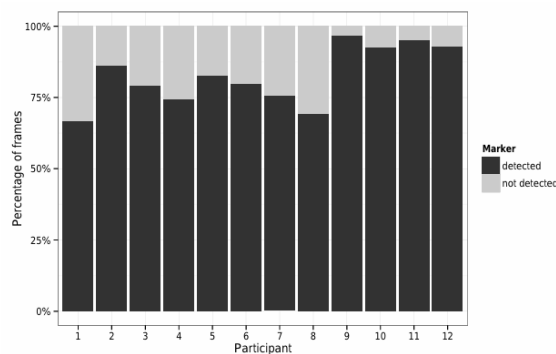


Рисунок 3.12. На рисунку показано відсоток кадрів, у яких положення системи відстеження погляду вдалося успішно визначити за допомогою відстеження маркерів методом «зсередини назовні».

Оцінка відстеження

Існує два ключові аспекти, що стосуються оцінки відстеження зсередини назовні на основі маркерів: маркери повинні бути виявлені на зображеннях, наданих камерою сцени, а положення системи відстеження очей має бути визначене на основі знайденого маркера з достатньою точністю.

У наступному розділі ми почнемо з аналізу того, скільки кадрів, записаних камерою сцени, ми можемо охопити, виявивши маркери. Якщо це неможливо, поточну позу системи відстеження очей неможливо визначити, і тому жоден промінь не може бути відкинутий у геометричну модель, а отже, - фіксацію неможливо класифікувати [35]. Далі за цим розділом йде розділ, присвячений аналізу точності оцінки пози.

Покриття

Орієнтацію та положення системи відстеження погляду можна оцінити для кожного кадру, в якому алгоритм комп'ютерного зору може ідентифікувати маркер. У 228150 кадрах наших тестових записів це стосувалося 186537 кадрів або 81,76 відсотка. Це означає, що 129,54 хвилини зі записаних 158,44 хвилин даних погляду можна класифікувати автоматично. На рисунку 4 показано деталі

для окремих учасників. Після перших восьми сеансів ми оптимізували розміщення та розмір маркерів позаду співрозмовника і таким чином змогли оптимізувати охоплення кадру до понад 92 відсотків. Те, що виявлення маркерів не вдалося в решті кадрів, має кілька причин. По-перше, іноді учасник дивився вбік, і маркера в полі зору просто не було. Однак частіше ці проблеми спричиняли швидкі рухи голови або різкі зміни положення. Результати, представлені тут, базуються на виявленні для кожного кадру та не включають відстеження пози від кадру до кадру, що може ще більше збільшити охоплення.

Ще один ефект, який можна спостерігати, полягає в тому, що чим довша фіксація, тим більша ймовірність виявлення маркерів під час поточної фіксації (див. рис. 3.13) .

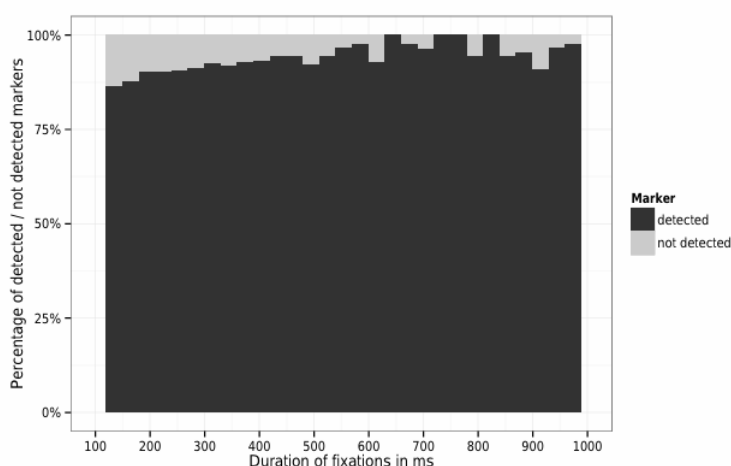


Рисунок 3.13 - На рисунку показано відсоток фіксацій (тривалістю понад 100 мс), для яких було виявлено маркер, тобто для яких вдалося визначити положення системи стеження за рухом очей, залежно від тривалості фіксацій

Під час стабільних фіксацій виявлення одного маркера призведе до успішної класифікації об'єкта інтересу. На рисунку 3.14 показано гістограму спостережуваних тривалостей, протягом яких не було виявлено жодного маркера. Близько 40 відсотків мають тривалість менше 100 мс і тому не є релевантними для багатьох аналізів значущих фіксацій.

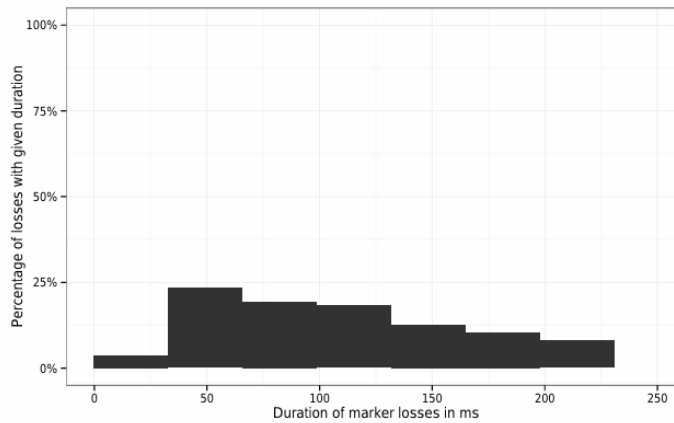


Рисунок 3.14 - Гістограма тривалості втрат маркерів, виявлених у ході експерименту. Медіана тривалості втрати маркерів становить 158,18 мс.

Точність

Системи відстеження «ззовні всередину» забезпечують точні результати з невеликою затримкою. З іншого боку, вони часто дорогі, обмежено портативні, і тому не можуть охопити великі площі. Відстеження «зсередини назовні», яке використовується в нашому підході, є дешевим і здатним охопити великі площі. Питання полягає в тому, чи достатньо високою є точність, якої можна досягти, для аналізу погляду.

Відстеження «ззовні всередину» проти відстеження «всередині назовні»

Щоб оцінити підхід відстеження «всередині назовні», в якому положення системи відстеження очей оцінюється шляхом виявлення та відстеження опорних маркерів, ми провели експеримент, використовуючи паралельно оптичну систему відстеження «ззовні всередину» DTrack2 від AR-Tracking GmbH з частотою дискретизації 60 Гц та точністю до міліметра. Таким чином, дані системи DTrack2 були взяті за еталонні дані.

Під час зіставлення оцінок пози, отриманих системою відстеження зовні всередину, з системою відстеження зсередини назовні, ми виявили, що остання мала середню затримку 379 мс (стандартне відхилення 90 мс), яка є стабільною протягом усього запису. Цю затримку можна пояснити часом, необхідним для обчислювально складнішого процесу.

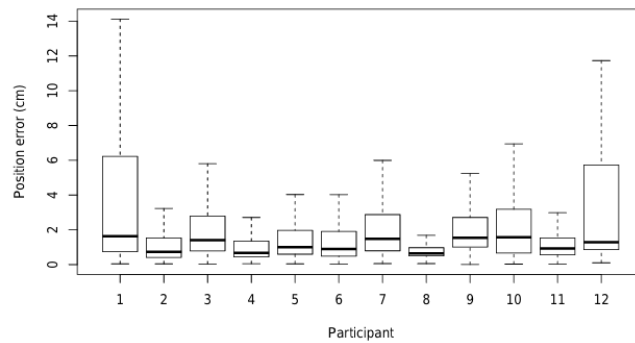


Рисунок 3.15 - Похибка положення для різних учасників при відстеженні «ззовні всередину» та «зсередини назовні».

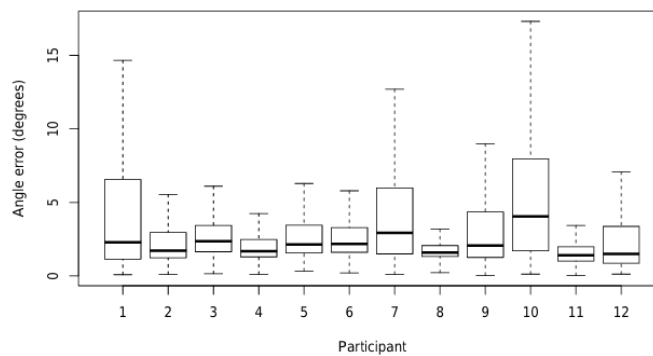


Рисунок 3.16 - Похибка кута для різних учасників при відстеженні «ззовні всередину» та «зсередини назовні».

дорогий ланцюжок обробки при використанні відео зі сцени-камери для оцінки пози: отримання відео, обчислення напрямку погляду, виявлення маркерів і, нарешті, фактична оцінка пози. Таким чином, більша затримка не є проблемою для онлайн- або офлайн-аналізу, оскільки система завжди правильно синхронізує оцінки погляду та положення. Однак це було б проблемою, якби використовувалася для інтерактивних застосувань, оскільки час реакції понад 400 мс є занадто високим для багатьох застосувань.

Вимірювання точності оцінки погляду вимагатиме вимірювання точності самого айтрекера [36]. У нашому дослідженні обидві системи відстеження пози використовувалися паралельно, тому будь-які помилки в оцінці напрямку погляду стосуються обох підходів. Тому тут ми розглядаємо лише точність оцінки пози.

Точність оцінки пози можна визначити з точки зору похибок положення (див. рис. 3.15) та кута (див. рис. 3.16) . Скоригована загальна середня похибка положення становить 1,23 см для трьох вимірів (стандартне відхилення 0,90 см). Скоригована середня похибка кута 2,25 градуса (стандартне відхилення 1,46 градуса) достатньо мала, щоб забезпечити точний розрахунок фіксацій на об'єктах, які не знаходяться в безпосередній близькості.

Помилки були скориговані шляхом видалення викидів, що є результатом технічного експерименту з двома додатковими маркерами, положення та орієнтація яких не були попередньо визначені в геометричній моделі, а вивчені онлайн під час їхньої першої зустрічі під час експериментів. Вони були розташовані на стіну за столом, щоб збільшити ймовірність відстеження, коли учасники дивилися вгору на свого співрозмовника. Ми тестували можливість почати з одного заздалегідь визначеного маркера та розширювати знання про геометрію на льоту [37]. Однак наш поточний підхід недостатньо стабільний і призводить до вищого стандартного відхилення, коли ці маркери є єдиними видимими, оскільки помилки під час початкової фази навчання ще не оптимізовані з часом.

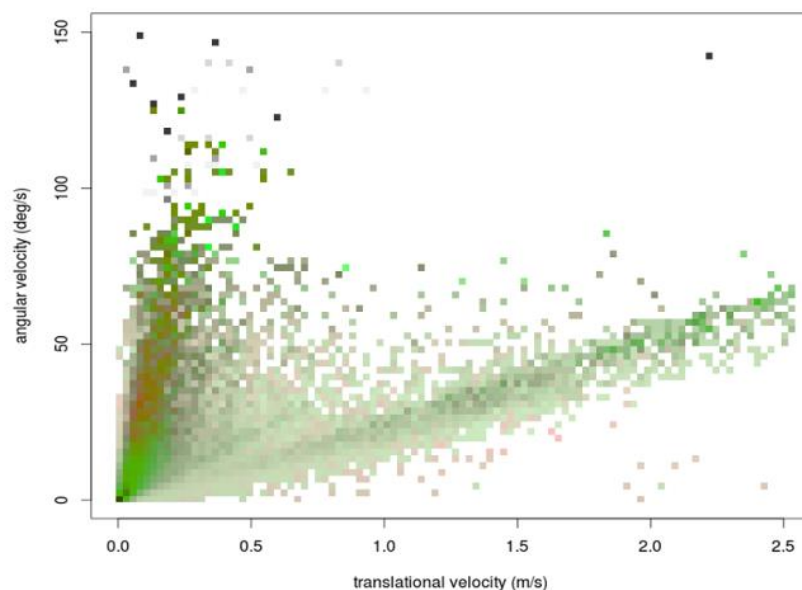


Рисунок 3.17 - Співвідношення виявлених і невиявлених маркерів залежно від кутової та поступальної швидкості.

Чим більше виявлених маркерів, тим зеленіший колір. Чим більше втрачених маркерів, тим червоніший колір. Інтенсивність кольору залежить від щільності їхнього розташування.

Хоча система відстеження зовні всередину є відносно стійкою до швидких рухів головки, вони призводять до розмиття в камері системи зсередини назовні. Тому, особливо для високих кутових швидкостей головки, швидкість виявлення маркерів зменшується. На рисунку 3.17 показано, що починаючи з кутових швидкостей понад 18 град/с, швидкість виявлення швидко зменшується (червона область). Чисто поступальні швидкості головки мають незначний вплив.

Оцінювання класифікації

Завдяки використанню підходу відстеження (всередині-назовні/зовні-всередину) положення трека очей та положення маркерів стають відомими. Для автоматичного анотування фіксацій на стимулах, розташування маркерів можна зробити явним, моделюючи їх у геометричній моделі стимулів як активні елементи.

Для описаної вище схеми експерименту людина-людина необхідними цільовими стимулами є фігури та обличчя співрозмовника. Фігури можуть бути частиною статичної геометричної моделі сценарію [38]. Таким чином, було створено 23 їх геометричні 3D-зображення. Оскільки нас цікавить лише те, чи дивилися на фігуру, ми могли опустити дрібні деталі. Таким чином, у цьому випадку для представлення фігур можна було б використовувати невеликі прямокутники, оскільки їх достатньо для наближеного відображення їхніх фактичних характеристик. Положення обличчя співрозмовника наближається прямокутником, розташованим у місці, де сидять учасники.

Після моделювання сценарію його необхідно вирівняти з реальною сценою, тобто з відео з камери сцени трека очей. Тому також необхідно змоделювати маркери, що використовуються для відстеження. Знаючи їх положення та орієнтацію в реальному світі, їхні віртуальні відповідники можна закріпити як накладання на сцену. Таким чином, у реальному світі створюється

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						69
Змн.	Арк.	№ докум.	Підпис	Дата		

ізоморфна система координат. Оскільки стимули та маркери знаходяться у фіксованих положеннях, положення стимулів залежать від маркерів. Таким чином, коли віртуальні маркери вирівняні з їхніми реальними відповідниками, проксі-геометрія стимулів також розташовується в правильному місці.

відомі положення маркерів у реальному світі, а також положення та орієнтація окулярів для відстеження погляду [39]. У поєднанні цього з напрямком погляду, виявленим системою відстеження погляду, перше можна зробити явним, змодельовавши його як промінь погляду [38]. Це дозволяє перевіряти зіткнення між проксі-геометрією стимулів та променем. Об'єкт інтересу фіксується, коли промінь погляду перетинається з проксі-геометрією, і, отже, лінія погляду перетинається зі стимулом у реальному світі.

Коли виявляється колізія, віртуальні проксі активно генерують подію, яка реєструється системою. Згенеровані події, а також усі відповідні дані із системи (вихідні дані відстеження очей, результати відстеження, додаткова інформація, така як позначки часу) записуються в розширений файл журналу, який використовується для постфактумного аналізу.

Ручна та автоматична класифікація

Ми маємо перші результати порівняння ручного анотування з автоматичною класифікацією. Для цього ми анотували відео одного учасника та порівняли дані з результатами, наданими EyeSee3D. Для 784 відповідних фіксацій автоматична класифікація збігалася з людською анотацією на 64,29 відсотка. Це добре, але ще не задовільний показник.

Це не можна пояснити лише втратою маркерів, оскільки у вибраному випадку коефіцієнт виявлення перевищує 90 відсотків. І навіть попри це, як показано на рисунку 6, маркери рідко втрачаються довше, ніж потрібно для фіксації інтересу (> 100 мс), тому реальна верхня межа має бути ще вищою.

На відміну від оцінки системи FixTag, два сеанси взаємодії тривалістю приблизно 2 хвилини кожен (184 та 172 фіксації відповідно) були автоматично проаналізовані з 99,5%/98,3% узгодженістю FixTag з ручними анотаторами.

Як можна пояснити різницю в продуктивності? По-перше, контекст

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						70
Змн.	Арк.	№ докум.	Підпис	Дата		

зовсім інший. Цільова сцена, що використовувалася для FixTag, складалася з 2D-площин, розташованих перпендикулярно, тому жодних перекриттів між різними областями інтересу не виникало. Крім того, цільові області в нашому дослідженні дуже малі, лише 3-4 см завширшки та 8 см заввишки, і вони щільно упаковані. Це призводить до багатьох випадків неоднозначностей під час анотування даних погляду. Як буде детально описано далі, це викликає більш загальні питання щодо класифікації даних погляду в 3D-сценаріях.

Досліджуючи випадки, коли автоматичне відстеження та анотації людиною розходилися між собою, ми виявили наступну основну тему: цільові об'єкти щільно упаковані. Таким чином, враховуючи один кадр з курсором погляду, часто важко вирішити, на якій фігурі насправді зосередився учасник. Саме з такою ситуацією зараз стикається автоматичний класифікатор, оскільки він обчислює перетин лінії зору для кожного кадру.

У цьому розділі представлено огляд галузі мобільної доповненої реальності (AR), включаючи історичний розвиток, майбутній потенціал, області застосування, виклики, компоненти та вимоги, найсучасніші системи та концепції інтерфейсу користувача. AR та носимі обчислення – це галузі, що швидко розвиваються, прикладом чого є стрімке зростання кількості дослідницьких внесків та комерційних розробок з середини 1990-х років. Досягли важливої точки в прогресі до мобільної доповненої реальності, оскільки доступна технологія є достатньо потужною для створення все більшої кількості вражаючих дослідницьких прототипів, але ще недостатньо надійною, універсальною та зручною для масового впровадження. Порівняно з іншими застосуваннями, описаними в цій книзі, які можна одразу реалізувати за допомогою сучасних технологій, мобільній доповненій реальності знадобиться більше часу, щоб досягти мейнстріму обчислювальної техніки. Однак мобільна доповнена реальність матиме величезний вплив, коли стане звичайною.

З нетерпінням чекаємо подальшого прогресу в галузях мініатюризації обчислювального обладнання, проектування акумуляторів, технології дисплеїв, сенсорної технології, точності та надійності відстеження, загального

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						71
Змн.	Арк.	№ докум.	Підпис	Дата		

відстеження на основі зору та розуміння сцени, а також загального комфорту. Ми очікуємо появи розподілених інфраструктур даних для контекстних обчислень загалом та доповненої реальності зокрема, що призведе до значного покращення можливостей доступу до даних для мобільних користувачів.

Зрештою, ми сподіваємося, що переваги мобільної доповненої реальності (AR) будуть досягнуті без шкоди для конфіденційності та комфорту. Дослідження та розробки в галузі, яка може суттєво вплинути на соціальні структури та звичаї, повинні супроводжуватися ретельним розглядом того, як можна захистити та зміцнити похвальні аспекти нашої соціальної рівноваги.

Незважаючи на зростаючий інтерес до доповненої реальності (AR) та великий обсяг досягнень і досліджень, все ще існує кілька викликів і питань, які потребують вирішення [40]. У цьому розділі ми класифікуємо обмеження, що характеризують сучасний стан розвитку AR, на основі таких аспектів: технології, соціальне сприйняття, зручність використання. У кожній з областей, описаних у цій роботі, досягнуто значного прогресу. Однак, технологія все ще має обмеження, які необхідно подолати. Система AR насправді має працювати з величезною кількістю інформації. Тому обладнання, яке використовується, має бути невеликим, легким, портативним і достатньо швидким для відображення графіки. Також термін служби батареї, що використовується цими складними пристроями AR, є ще одним обмеженням для використання AR. Крім того, AR-відстеження потребує певного системного обладнання, такого як GPS, для забезпечення точного маркера, що вимагає від нього як точності, так і достатньої надійності. Ці апаратні перешкоди необхідно вирішити для практичного використання AR. Системи AR зазвичай отримують багато інформації та потребують програмного забезпечення для фільтрації інформації, збереження корисної інформації, відкидання непотрібних даних та її зручного відображення.

Існує кілька можливих напрямків майбутніх досліджень. Багато шлемів для моніторингу (HMD), створених спеціально для доповненої реальності (AR), потребують розробки. HMD все ще занадто громіздкі та мають обмежене поле зору, контрастність та роздільну здатність. HMD та інше носимне обладнання,

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						72
Змн.	Арк.	№ докум.	Підпис	Дата		

таке як рукавички для обробки даних та костюми для обробки даних, є обмеженням для користувача. Усе носимне обладнання має бути розроблене таким чином, щоб воно було легшим, меншим та простішим у використанні для користувача. Також дослідникам систем доповненої реальності необхідно враховувати інші проблеми, такі як затримки часу відгуку, апаратні або програмні збої систем доповненої реальності. Одним з обмежень систем доповненої реальності є помилка реєстрації. Виявлення оклюзії є активною областю дослідження систем доповненої реальності. Аналізуючи різні методи відстеження, визначаються можливі напрямки досліджень відстеження, які дозволяють дослідникам ефективно використовувати знання у відеокдрах або інтегрувати методи на основі зору з іншими датчиками.

3.4 Висновок по розділу

У третьому розділі було розглянуто розробку та інтеграцію комп'ютерного зору в мобільний AR-додаток. Проведено проєктування архітектури мобільного застосунку з використанням технологій доповненої реальності та комп'ютерного зору. Реалізовано програмну систему розпізнавання та відстеження об'єктів у AR-середовищі, що забезпечує інтерактивну взаємодію користувача з цифровими елементами. Також виконано тестування та оцінку ефективності алгоритмів відстеження й розпізнавання, що дозволило оцінити стабільність і точність роботи системи. Отримані результати підтверджують ефективність інтеграції комп'ютерного зору в мобільні AR-додатки для підвищення функціональності та якості взаємодії користувача з доповненою реальністю.

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						73
Змн.	Арк.	№ докум.	Підпис	Дата		

ВИСНОВКИ

У ході виконання бакалаврської роботи було досліджено процес інтеграції комп'ютерного зору в мобільні додатки доповненої реальності та розроблено програмне рішення, що демонструє практичне застосування сучасних методів обробки зображень, розпізнавання та відстеження об'єктів у реальному часі. Виконана робота є комплексним процесом, спрямованим на створення ефективного, продуктивного та адаптивного мобільного AR-додатку, здатного працювати в умовах обмежених ресурсів мобільних пристроїв.

У межах дослідження було проаналізовано теоретичні основи комп'ютерного зору та його роль у системах доповненої реальності, розглянуто архітектуру мобільних AR-додатків, а також визначено основні проблеми їх реалізації, зокрема точність відстеження, вплив умов освітлення та обмеження обчислювальних ресурсів. Це дозволило сформулювати розуміння ключових викликів, які виникають при інтеграції комп'ютерного зору в AR-середовище.

Далі було досліджено сучасні методи комп'ютерного зору, включаючи алгоритми обробки зображень, детекції та трекінгу об'єктів, а також порівняно підходи доповненої та віртуальної реальності. На основі цього було обґрунтовано вибір інструментів і технологій для реалізації мобільного AR-додатку, що забезпечують стабільну роботу системи та ефективну обробку візуальних даних у реальному часі.

На етапі практичної реалізації було розроблено архітектуру мобільного додатку з інтегрованим комп'ютерним зором, реалізовано основні модулі обробки зображень і відстеження об'єктів, а також забезпечено взаємодію між компонентами системи. Особливу увагу було приділено оптимізації продуктивності, оскільки мобільні пристрої мають обмежені ресурси для виконання складних обчислень у реальному часі.

Було проведено тестування розробленого рішення, під час якого оцінювалася точність розпізнавання, стабільність відстеження та швидкість обробки даних. Результати тестування показали, що система демонструє

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						74
Змн.	Арк.	№ докум.	Підпис	Дата		

задовільну продуктивність у реальних умовах використання, однак у складних сценах або при змінному освітленні можливе зниження точності, що потребує подальшої оптимізації алгоритмів.

У результаті виконання роботи було підтверджено, що інтеграція комп'ютерного зору в мобільні AR-додатки значно розширює їх функціональні можливості, забезпечує більш точну взаємодію з реальним середовищем і підвищує рівень занурення користувача. Використання сучасних методів обробки зображень дозволяє створювати більш інтелектуальні та адаптивні системи доповненої реальності.

Загалом розроблений мобільний AR-додаток відповідає поставленій меті та вимогам дослідження. До його переваг можна віднести ефективну інтеграцію комп'ютерного зору, можливість роботи в реальному часі та розширювану архітектуру. Водночас існують певні обмеження, пов'язані з продуктивністю на слабких пристроях та чутливістю до умов середовища.

У перспективі подальших досліджень можливим є впровадження більш складних алгоритмів глибокого навчання для підвищення точності розпізнавання, оптимізація енергоспоживання мобільних пристроїв, а також розширення функціональності системи за рахунок підтримки багатокористувацьких AR-сценаріїв та інтеграції з хмарними сервісами.

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
						75
Змн.	Арк.	№ докум.	Підпис	Дата		

СПИСОК ПОСИЛАНЬ НА ДЖЕРЕЛА

1. OpenCV Documentation. (2025). <https://docs.opencv.org/>
2. Unity AR Foundation. (2025). <https://docs.unity3d.com/Packages/com.unity.xr.arfoundation>
3. Apple Developer ARKit Documentation. (2025). <https://developer.apple.com/augmented-reality/arkit/>
4. Google ARCore Documentation. (2025). <https://developers.google.com/ar>
5. Microsoft Mixed Reality Toolkit. (2025). <https://learn.microsoft.com/en-us/windows/mixed-reality/mrtk-unity/>
6. Szeliski, R. (2022). Computer Vision: Algorithms and Applications. Springer. <https://szeliski.org/Book/>
7. Forsyth, D., & Ponce, J. (2018). Computer Vision: A Modern Approach. Pearson.
8. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press. <https://www.deeplearningbook.org/>
9. Bradski, G. (2024). Learning OpenCV 4. O'Reilly Media.
10. Zhang, Z. (2023). Camera calibration in computer vision. IEEE Transactions on Pattern Analysis. <https://doi.org/10.1109/TPAMI>
11. Lowe, D. G. (2004). Distinctive Image Features from Scale-Invariant Keypoints. IJCV.
12. Viola, P., & Jones, M. (2001). Rapid Object Detection using a Boosted Cascade. CVPR.
13. Redmon, J. (2023). YOLO: Real-Time Object Detection. <https://pjreddie.com/darknet/yolo/>
14. He, K., Zhang, X., Ren, S. (2016). Deep Residual Learning for Image Recognition. CVPR.
15. Szegedy, C. (2015). Going deeper with convolutions. CVPR.

					Б.ІІІ - 82.00.00.000 ІІЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		76

16. Unity Technologies. (2025). AR Development Best Practices.
<https://unity.com/solutions/ar>
17. Apple Vision Framework Documentation. (2025).
<https://developer.apple.com/documentation/vision>
18. Google ML Kit for Mobile Vision. (2025).
<https://developers.google.com/ml-kit>
19. TensorFlow Lite for Mobile. (2025). <https://www.tensorflow.org/lite>
20. PyTorch Mobile Documentation. (2025). <https://pytorch.org/mobile/>
21. OpenCV AI Kit (OAK). (2024). <https://opencv.ai/>
22. Kato, H., & Billinghurst, M. (1999). Marker Tracking and HMD Calibration for AR.
23. Azuma, R. (1997). A Survey of Augmented Reality. Presence Journal.
24. Billinghurst, M., Clark, A., & Lee, G. (2015). A Survey of Augmented Reality. Foundations and Trends in HCI.
25. Craig, A. (2013). Understanding Augmented Reality. Morgan Kaufmann.
26. Rekimoto, J. (2022). SmartAR: Augmented Reality Interaction. ACM Digital Library.
27. Feiner, S. (2002). Augmented Reality: A New Way of Seeing. IEEE Computer.
28. Khronos Group. (2025). OpenXR Specification.
<https://www.khronos.org/openxr/>
29. Apple RealityKit Documentation. (2025).
<https://developer.apple.com/documentation/realitykit>
30. Google MediaPipe Framework. (2025).
<https://developers.google.com/mediapipe>
31. Shi, J., & Tomasi, C. (1994). Good Features to Track. CVPR.
32. Lucas, B. D., & Kanade, T. (1981). An Iterative Image Registration Technique.
33. Hartley, R., & Zisserman, A. (2004). Multiple View Geometry in Computer Vision. Cambridge University Press.

					Б.ІІІ - 82.00.00.000 ІІЗ	Арк.
						77
Змн.	Арк.	№ докум.	Підпис	Дата		

34. Nvidia DeepStream SDK. (2025).
<https://developer.nvidia.com/deepstream-sdk>
35. Intel OpenVINO Toolkit. (2025). <https://docs.openvino.ai/>
36. AWS Panorama for Computer Vision. (2025).
<https://aws.amazon.com/panorama/>
37. Azure AI Vision Services. (2025). <https://learn.microsoft.com/en-us/azure/ai-services/computer-vision/>
38. Li, S., & Lee, D. (2024). Mobile Augmented Reality Systems: Design and Challenges. IEEE Access.
39. Wang, Q., & Chen, Y. (2023). Real-Time Visual Tracking in Mobile AR Applications. ACM Computing Surveys.
40. Zhang, H., & Liu, M. (2024). Edge AI for Mobile Computer Vision Applications. Journal of Systems Architecture.

БІБЛІОГРАФІЧНА ДОВІДКА

Тема бакалаврської роботи: " Інтеграція комп'ютерного зору в мобільні додатки доповненої реальності "

Обсяг пояснювальної записки: 75 аркушів

Дата закінчення бакалаврської роботи 10 червня 2026р.

Підпис студента _____

					Б.ПІ - 82.00.00.000 ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		79