

МАГІСТЕРСЬКА РОБОТА

МР. ШМ - 45.00.00.000 ПЗ

Група ШМ-24-3

Павликівський Андрій

2025

Івано-Франківський національний технічний університет нафти і газу

Факультет інформаційних технологій

Кафедра інженерії програмного забезпечення

Павликівський Андрій Васильович

(прізвище, ім'я, по батькові)

УДК 004.9
(індекс)

МАГІСТЕРСЬКА РОБОТА

Методи та моделі масштабування мультимодальних інтерфейсів у

віртуальній та доповненій реальності

(назва роботи)

Інженерія програмного забезпечення

(назва освітньої програми)

121 - Інженерія програмного забезпечення

(шифр і назва спеціальності)

Павликівський А.В.

(підпис, ініціали та прізвище здобувача освітнього ступеня)

Науковий керівник Юрчишин Володимир Миколайович, д.т.н., професор

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

Допущено до захисту

Завідувач кафедри

доц. Бандура В.В.

(посада) (підпис) (дата) (ініціали та прізвище)

Нормоконтроль

доц. Вовк Р.Б.

(посада) (підпис) (дата) (ініціали та прізвище)

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

Івано-Франківськ – 2025

Івано-Франківський національний технічний університет нафти і газу

Факультет інформаційних технологій

Кафедра інженерії програмного забезпечення

Освітній рівень магістр

Спеціальність 121 – Інженерія програмного забезпечення

ЗАТВЕРДЖУЮ:

Зав. кафедрою

ІІЗ

доц.

В.В. Бандура

“ 04 ” вересня 2025 р.

ЗАВДАННЯ

НА МАГІСТЕРСЬКУ РОБОТУ СТУДЕНТУ

Павликівському Андрію Васильовичу

(прізвище, ім'я, по-батькові)

1. Тема магістерської роботи “**Методи та моделі масштабування мультимодальних інтерфейсів у віртуальній та доповненій реальності**”

керівник проекту (роботи) Юрчишин В.М., д.т.н., професор

затверджені наказом закладу вищої освіти від “ 05 ” листопада 2025 р. № 695/7

2. Строк подання студентом проекту (роботи) 15 грудня 2025 р.

3. Вихідні дані до проекту (роботи) Формальні моделі, концепції та методи побудови інформаційних та програмних технологій масштабування мультимодальних інтерфейсів

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

1. Дослідження проблеми масштабування інтерфейсів у віртуальній та доповненій реальності

2. Методологія проектування інтерфейсу за допомогою швидкого прототипування для VR

3. Методологія проектування віртуальних агентів у середовищах змішаної реальності

4. Реалізація моделей та методів масштабування мультимодальних інтерфейсів у VR та AR

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

1. Континуум реальність-віртуальність (рис. 1.1)

2. Вигляд ігрового рушія Unity 3D (рис. 1.2)

3. Система візуального скриптингу Blueprints (рис. 1.3)

4. Приклад Reality Composer програми (рис. 1.4)

5. Приклад роботи Wired Actions (рис. 1.5)

6. Консультанти розділів проекту (роботи)

Розділ	Консультант	Підпис, дата
Перевірка на плагіат	доц., к.т.н. Вовк Р.Б.	

7. Дата видачі завдання 04 вересня 2025 р.

Керівник

_____ (підпис)

Завдання прийняв до виконання _____

(підпис)

КАЛЕНДАРНИЙ ПЛАН

№ п/п	Назви етапів магістерської роботи	Строк виконання етапів роботи	Примітка
1	Підбір і вивчення літератури по темі магістерської роботи	15.09.2025	виконано
2	Дослідження проблеми масштабування інтерфейсів у віртуальній та доповненій реальності	01.10.2025	виконано
3	Методологія проектування інтерфейсу за допомогою швидкого прототипування для VR	17.10.2025	виконано
4	Методологія проектування віртуальних агентів у середовищах змішаної реальності	02.11.2025	виконано
5	Реалізація моделей та методів масштабування мультимодальних інтерфейсів у VR та AR	19.11.2025	виконано
6	Архітектура та технологічна платформа системи змішаної реальності	02.12.2025	виконано
7	Затвердження пояснювальної записки роботи завідувачем кафедри	15.12.2025	виконано

Студент – магістр

_____ (підпис)

Керівник роботи

_____ (підпис)

АНОТАЦІЯ

Магістерська робота: 76 с., 26 рис., 11 табл., 41 джерело.

Тема: Методи та моделі масштабування мультимодальних інтерфейсів у віртуальній та доповненій реальності

Метою роботи є розроблення та обґрунтування методів і моделей масштабування мультимодальних інтерфейсів у віртуальній та доповненій реальності шляхом створення методології реалізації XR-системи з підтримкою природної взаємодії.

Об'єкт дослідження - процеси побудови, інтеграції та масштабування мультимодальних інтерфейсів у системах віртуальної та доповненої реальності.

Предмет дослідження - методи, моделі, архітектурні підходи та технологічні інструменти, що забезпечують масштабованість та ефективність мультимодальної взаємодії в XR-середовищах.

Результати дослідження

В роботі описано механізми підтримання масштабованості агентної системи, що дозволяють розширювати її функціональність без повної зміни архітектурного ядра та забезпечують стабільність при різних рівнях навантаження.

Висновок

Було реалізовано архітектурні рішення для багатомодальної взаємодії, включно з моделлю діалогового потоку, розпізнаванням об'єктів поглядом та впровадженням інтелектуального віртуального агента.

МУЛЬТИМОДАЛЬНА ВЗАЄМОДІЯ, ВІРТУАЛЬНА РЕАЛЬНОСТІ, ДОПОВНЕНА РЕАЛЬНОСТІ, МАСШТАБУВАННЯ ІНТЕРФЕЙСІВ, ПРОТОТИПУВАННЯ, ІНТЕЛЕКТУАЛЬНІ АГЕНТИ, АРХІТЕКТУРА XR-СИСТЕМ, ДІАЛОГОВІ МОДЕЛІ, UX-ДИЗАЙН.

ABSTRACT

Master Thesis: 76 pp., 26 fig., 11 tab., 41 sources.

Topic: Methods and models for scaling multimodal interfaces in virtual and augmented reality

The purpose of the work is to develop and substantiate methods and models for scaling multimodal interfaces in virtual and augmented reality by creating a methodology for implementing an XR system with support for natural interaction.

The object of the study is the processes of building, integrating and scaling multimodal interfaces in virtual and augmented reality systems.

The subject of the study is methods, models, architectural approaches and technological tools that ensure scalability and efficiency of multimodal interaction in XR environments.

Research results

The paper describes mechanisms for maintaining the scalability of the agent system, which allow expanding its functionality without completely changing the architectural core and ensure stability at different load levels.

Conclusion

Architectural solutions for multimodal interaction were implemented, including a dialog flow model, object recognition by gaze and the implementation of an intelligent virtual agent.

MULTIMODAL INTERACTION, VIRTUAL REALITY, AUGMENTED REALITY, SCALING INTERFACES, PROTOTYPING, INTELLIGENT AGENTS, XR SYSTEM ARCHITECTURE, DIALOGUE MODELS, UX DESIGN.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ	10
ВСТУП.....	11
РОЗДІЛ 1. ДОСЛІДЖЕННЯ ПРЕДМЕТНОЇ ОБЛАСТІ МАСШТАБУВАННЯ ІНТЕРФЕЙСІВ У ВІРТУАЛЬНІЙ ТА ДОПОВНЕНІЙ РЕАЛЬНОСТІ.....	15
1.1. Дослідження масштабованості інтерфейсів користувача в крос-платформенних додатках розширеної реальності.....	15
1.2. Континуум віртуальності та класифікація технологій розширеної реальності.....	18
1.3. Роль швидкого прототипування в проектуванні інтерфейсів розширеної реальності.....	20
1.3.1. Переваги та специфіка інтерактивного прототипування.....	20
1.3.2. Проблема дистанційного тестування та вибір методології.....	21
1.4. Дослідження існуючих інструментів швидкого прототипування для XR	22
1.4.1. Інструменти прототипування на основі скриптингу/ігрових двигунів (Expert Level)	22
1.4.2. Інструменти для візуального скриптингу та «No-Code» / «Low-Code» (Intermediate Level)	24
1.4.3. Інструменти на Основі 2D-Макетів та Імітації (Conceptual Level) .	25
1.4.4. Нові інструменти для швидкого моделювання XR.....	26
1.5. Концептуалізація та дослідницькі прогалини масштабованих крос-реальних систем	27
1.6. Розробка принципів дизайну інтерфейсів користувача для розширеної реальності.....	28
1.6.1. Специфічні виклики UI в XR-системах.....	29
1.6.2. Існуючі спеціалізовані рекомендації	29
Висновки до розділу	30

РОЗДІЛ 2. МЕТОДОЛОГІЯ ПРОЕКТУВАННЯ ІНТЕРФЕЙСУ ЗА ДОПОМОГОЮ ШВИДКОГО ПРОТОТИПУВАННЯ ДЛЯ ВІРТУАЛЬНОЇ РЕАЛЬНОСТІ.....	32
2.1. Концептуалізація персонажів користувачів та афініті діаграмування..	32
2.2. Основні етапи проектування початкових ескізів інтерфейсу та віртуального дизайну.....	34
2.2.1. Проектування ескізів	34
2.2.2. Визначення завдань для прототипування макетів	36
2.2.3. Прототипування засобами Adobe XD.....	37
2.3. Методологія проектування інтелектуальних віртуальних агентів у середовищах змішаної реальності	43
Висновки до розділу	45
РОЗДІЛ 3. РЕАЛІЗАЦІЯ МОДЕЛЕЙ ТА МЕТОДІВ МАСШТАБУВАННЯ МУЛЬТИМОДАЛЬНИХ ІНТЕРФЕЙСІВ У ВІРТУАЛЬНІЙ ТА ДОПОВНЕНІЙ РЕАЛЬНОСТІ	47
3.1. Імплементация дизайну інтерфейсу в мультимодальну архітектуру візуалізації.....	47
3.1.1. Аналіз та декомпозиція існуючої архітектури	47
3.1.2. Інтеграція прототипу та модифікація UI	49
3.2. Архітектурні вимоги та методи багатомодальної взаємодії.....	49
3.2.1. Багатомодальна взаємодія. Архітектура потоку діалогової взаємодії	53
3.2.2. Розпізнавання об'єктів поглядом	56
3.2.3. Архітектура інтелектуального віртуального агента	56
3.3. Візуальна експресія та анімаційний контроль віртуального персонажа.....	58
3.3.1. Архітектура синхронізації анімацій.....	59
3.3.2. Архітектура обробки запиту та генерації відповіді віртуального персонажа.....	61

3.4. Забезпечення масштабованості та роботи віртуального агента	63
3.5. Архітектура та технологічна платформа системи змішаної реальності	64
Висновки до розділу	67
ВИСНОВКИ	69
ПЕРЕЛІК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ.....	72

ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ

MR - Mixed Reality

VR - Virtual Reality

AR - Augmented Reality

XR - Extended Reality

UI - User Interface

VIMS - Visually Induced Motion Sickness

SME - Subject Matter Expert

TTS - Text-to-Speech

STT - Speech-to-Text

CR - Cross-Reality

HMD - Head-Mounted Display

FAB - Floating Action Button

MRTK - Mixed Reality Toolkit

GED - General Educational Development

ВСТУП

Актуальність теми.

Стрімкий розвиток технологій віртуальної (VR), доповненої (AR) та змішаної (MR) реальності призвів до формування нового класу інтерактивних систем, здатних забезпечувати глибоку мультимодальну взаємодію між користувачем та цифровим середовищем. Зі збільшенням доступності XR-пристроїв та розширенням сфер їх застосування виникає потреба у створенні інтерфейсів, які можуть ефективно масштабуватися, адаптуватися до різних апаратних платформ і підтримувати природну взаємодію на основі жестів, голосу, погляду та просторових маніпуляцій. Особливого значення набувають методи швидкого прототипування, що дозволяють прискорити розробку та експериментальне тестування XR-інтерфейсів, забезпечуючи перевірку дизайну до етапу впровадження.

Незважаючи на значні досягнення у сфері XR, процес створення мультимодальних інтерфейсів залишається складним і слабо формалізованим. Наявні інструменти та підходи демонструють фрагментарність, що ускладнює побудову універсальних моделей масштабування. Сучасні системи потребують архітектурних рішень, здатних підтримувати когнітивні агенти, адаптивні компоненти UI та високоефективні методи обробки мультимодальних сигналів. Тому розробка системного підходу до моделювання та масштабування інтерфейсів у XR-середовищах має важливе наукове та практичне значення.

У роботі запропоновано комплексне дослідження архітектур, методологій проектування та засобів реалізації мультимодальних XR-інтерфейсів, спрямоване на створення узгодженої моделі масштабування. Отримані результати дозволяють не лише визначити ключові принципи побудови інтерфейсів, але й забезпечити можливість розширення їх функціональності відповідно до потреб користувачів та обмежень технічних платформ.

Актуальність дослідження зумовлена зростаючим попитом на XR-рішення у сферах освіти, медицини, індустрії 4.0, культурної спадщини, тренажерних систем та цифрових сервісів нового покоління. Збільшення кількості пристроїв з підтримкою мультимодальної взаємодії створює нові можливості, але водночас підвищує вимоги до масштабованості інтерфейсів. Сучасні системи часто не мають уніфікованих підходів до адаптації UI під різні формати взаємодії: від ширококутних AR-окулярів до повноцінних VR-шоломів, що створює труднощі у забезпеченні стабільної та ефективної роботи. Крім того, відсутність стандартизованих моделей проєктування мультимодальних інтерфейсів ускладнює реалізацію когнітивних агентів, які мають реагувати на природні дії користувача — погляд, інтонацію, рухи тіла. Ці виклики вимагають створення нових методів проєктування, що враховують контекст XR-простору, обмеження апаратних платформ, поведінкові особливості користувача та засоби дистанційного тестування.

У цьому контексті розробка моделей масштабування XR-інтерфейсів стає критичною для побудови систем, здатних до гнучкого розширення, повторного використання компонентів та швидкого впровадження інтелектуальних агентів. Саме тому запропоноване дослідження є важливим внеском у формування методологічної основи для розвитку сучасних XR-технологій.

Метою роботи є розроблення та обґрунтування методів і моделей масштабування мультимодальних інтерфейсів у віртуальній та доповненій реальності шляхом створення методології реалізації XR-системи з підтримкою природної взаємодії.

Об'єкт дослідження - процеси побудови, інтеграції та масштабування мультимодальних інтерфейсів у системах віртуальної та доповненої реальності.

Предмет дослідження - методи, моделі, архітектурні підходи та технологічні інструменти, що забезпечують масштабованість та ефективність мультимодальної взаємодії в XR-середовищах.

Завдання дослідження:

1. Проаналізувати сучасні моделі та інструменти масштабування XR-інтерфейсів.
2. Дослідити континуум віртуальності та класифікувати технології XR у контексті вимог до UI/UX.
3. Вивчити методи швидкого прототипування та визначити їх застосовність для XR-дизайну.
4. Розробити методологію проектування XR-інтерфейсів із використанням мультимодальної взаємодії.
5. Побудувати архітектуру інтелектуального віртуального агента.
6. Реалізувати прототип XR-системи з підтримкою мультимодальної взаємодії.

Методи дослідження

У роботі використано комплекс методів, зокрема:

- системний аналіз для визначення структурних компонентів XR-систем;
- методи UX-досліджень (персони, афініті-діаграмування, прототипування);
- моделювання архітектурних структур мультимодальної взаємодії;
- методи проектування діалогових агентів;
- алгоритмічні підходи до обробки жестів, голосу та погляду;
- порівняльний аналіз інструментів і платформ XR.

Наукова новизна отриманих результатів

Наукова новизна полягає у створенні комплексної моделі масштабування мультимодальних XR-інтерфейсів, яка поєднує архітектурні принципи, методи UX-проективання та алгоритми мультимодальної взаємодії. У роботі вперше запропоновано узгоджений підхід до побудови інтелектуального агента в XR, що інтегрує голосову, жестову та поглядову взаємодію у єдину когнітивно узгоджену систему. Розроблена архітектура

синхронізації анімацій забезпечує природність комунікації між користувачем і віртуальним персонажем.

Практичне застосування результатів

Результати роботи можуть бути використані у процесі розроблення навчальних, медичних, тренажерних, промислових та комунікаційних XR-систем. Створені моделі та архітектурні рішення можуть бути інтегровані в ігрові рушії, системи симуляції, платформи змішаної реальності. Методологія проєктування може застосовуватися командами розробників XR-додатків для прискорення роботи над інтерфейсами та агентовими підсистемами.

Структура магістерської роботи. Представлена робота складається зі вступу, трьох розділів та висновків. Загальний обсяг роботи становить 76 сторінок, і містить 26 рисунків, 11 таблиць, перелік використаних джерел із 41 найменування.

РОЗДІЛ 1. ДОСЛІДЖЕННЯ ПРЕДМЕТНОЇ ОБЛАСТІ МАСШТАБУВАННЯ ІНТЕРФЕЙСІВ У ВІРТУАЛЬНІЙ ТА ДОПОВНЕНІЙ РЕАЛЬНОСТІ

1.1. Дослідження масштабованості інтерфейсів користувача в крос- платформенних додатках розширеної реальності

Застосування технологій розширеної реальності (XR) стимулювало розширення нових напрямків досліджень. Термін XR часто використовується як загальна категорія, що охоплює віртуальну реальність (VR), доповнену реальність (AR) та змішану реальність (MR), концептуалізованих у межах континууму реальність-віртуальність Мілграма [2].

Цей континуум являє собою безперервну шкалу, на якій можна розмістити будь-яку систему на основі її ступеня занурення та співвідношення фізичного й цифрового контенту.

Континуум візуально демонструє, що розширена реальність (Extended Reality, XR) є загальним терміном, що охоплює всі точки на цій шкалі (VR, AR, MR).

Хоча континуум Мілграма є концептуальною моделлю, а не алгоритмом, що виконується, його можна пояснити через основні принципи технологічної реалізації для кожної точки.

Таблиця 1.1.

Технологічні особливості континууму Мілграма

Категорія	Ключовий принцип роботи	Основні вимоги до обробки	Приклади апаратного забезпечення
VR (віртуальна реальність)	Повне генерування цифрового світу та блокування фізичного зору.	Висока продуктивність графіки (рендеринг 3D-сцени) та відстеження руху голови для мінімізації затримки.	Oculus Quest, HTC Vive, Sony PS VR.

Категорія	Ключовий принцип роботи	Основні вимоги до обробки	Приклади апаратного забезпечення
AR (доповнена реальність)	Позиційне відстеження реального світу та накладання віртуальних об'єктів на реальний відеопотік або прозоре скло.	Одночасна обробка: 1) SLAM (Simultaneous Localization and Mapping) для визначення положення користувача в реальному просторі. 2) Рендеринг віртуальних об'єктів.	Мобільні телефони (ARKit/ARCore), Google Glass, Microsoft HoloLens.
MR (змішана реальність)	Глибока інтеграція реального та віртуального, що дозволяє віртуальним об'єктам взаємодіяти з реальним середовищем (наприклад, віртуальний об'єкт може бути закритий реальним).	Висока точність SLAM, розуміння геометрії реального середовища, просторове закріплення віртуального контенту та його маскуванню реальністю (Occlusion).	Microsoft HoloLens 2, Magic Leap One.

Таким чином, "алгоритм роботи" в даному контексті зводиться до складності імплементації та балансу між реальним і віртуальним контентом, що зростає в міру руху до центру континууму (MR), де необхідне високоточне злиття обох світів.

Останнім часом спостерігається значне зростання впровадження та популярності XR, що зумовлено прогресом мобільних технологій та розробкою споживчих шоломів з дисплеями (HMD). Очікується, що ця тенденція збережеться, оскільки прогнозується, що ринки технологій XR досягнуть понад 2 мільярди доларів до 2027 року [5]. В міру подальшого розвитку цієї технології ряд галузей виявив інтерес до інтеграції XR для інноваційних застосувань у сферах охорони здоров'я, освіти, навчання, виробництва та інших [6].

Зростаючий попит на технології XR також сприяв підвищенню доступності систем, оскільки на споживчий ринок виходять більш економічні варіанти. Як наслідок, використання додатків XR стало більш поширеним у повсякденному житті, прикладом чого є мобільні смартфони з можливостями

AR [7]. Однак розробка додатків XR, особливо тих, що функціонують у кількох модальностях, все ще перебуває на початкових етапах. Крім того, розробка стандартів, умовностей та найкращих практик для дизайну інтерфейсу користувача (UI) XR загалом вважається такою, що потребує додаткових досліджень [8]. Хоча були проведені деякі дослідження існуючих середовищ XR, зростаюча кількість додатків, які використовують більше одного типу реальності, призводить до дефіциту досліджень, що стосуються дизайну інтерфейсів користувача, спеціально розроблених для цих крос-реальних додатків. Поточні інтерфейси часто базуються на обмеженнях апаратного та програмного забезпечення, зручності або швидкості розробки, з низьким пріоритетом на застосуванні теоретичних засад при проектуванні.

Метою цього дослідження є вивчення створення інтерфейсів користувача в крос-платформених додатках XR з кінцевою метою розробки більш масштабованих та багаторазових інтерфейсів користувача в різних реаліях та на різних апаратних засобах. У цьому контексті масштабованість визначається як здатність керувати інформацією інтерфейсу користувача в різних режимах з огляду на розвиток технологій.

Це дослідження має на меті ідентифікувати фактори інтерфейсу користувача, які є спільними для різних систем, для розробки керівних принципів дизайну, адаптованих для багатомодальних інтерфейсів користувача. Крім того, завданням є вдосконалення поточних інтерфейсів, щоб дисплеї були зрозумілими, зручними та оптимальними для успішного планування. Вдосконалення UI включають дослідження кількості представленої інформації, типів інформації, представлення піктограм, ієрархій меню та інших елементів. Це дослідження також досліджує сприйнятту складність реалізованих рішень дизайну UI, щоб гарантувати, що обсяг інформації, представленої користувачам, дозволяє здійснювати планування місій протягом розумного періоду часу.

Для дослідження цих факторів було проведено деталізацію дослідження, а також необхідний дизайн та реалізація інтерфейсів

користувача. У цій роботі використовується модель ітеративного дизайну, побудови та оцінки, яка є центральною для взаємодії людини з комп'ютером (HCI) та дизайну, орієнтованого на користувача.

1.2. Континуум віртуальності та класифікація технологій розширеної реальності

Концептуальна модель континууму реальність-віртуальність (Virtuality Continuum), вперше представлена у [5], є фундаментальною структурою для класифікації технологій, які інтегрують реальні та віртуальні світи на основі ступеня віртуальності. Цей континуум розміщує реальне середовище на одному (лівому) краю, а повністю віртуальне середовище — на протилежному (правому) кінці. Рух вздовж континууму у правий бік означає збільшення ступеня віртуальності та зменшення безпосередньої взаємодії з фізичним світом.

Доповнена реальність (AR) займає те місце в континуумі, де віртуальні об'єкти занурюються у переважно реальне світове середовище (рис. 1.1).

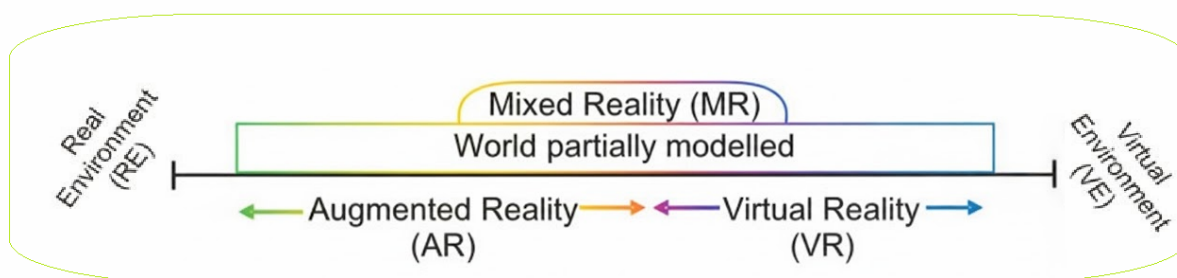


Рис. 1.1. Континуум реальність-віртуальність

Для потреб цього дослідження, нижче наведено детальний опис класифікації реалій, які обговорюються в подальшому тексті:

1. Віртуальна Реальність (Virtual Reality, VR)

Віртуальна реальність (VR) розміщується на крайньому правому кінці континууму і являє собою повне занурення користувача у комп'ютерно

згенероване, синтетичне середовище, де фізичний реальний світ повністю блокується. Це досягається завдяки використанню шоломів з дисплеями (HMDs), які надають високоіммерсивні, візуальні та тривимірні симуляційні характеристики. Ключова особливість VR полягає у відсутності взаємодії з реальним світом, що дозволяє створювати повністю контрольовані симуляційні сценарії, широко застосовувані для тренувань, планування та розваг.

2. Змішана Реальність (Mixed Reality, MR)

Змішана реальність (MR) охоплює весь середній простір континууму, розташований між чистим реальним та чистим віртуальним середовищами. Вона визначається як злиття реального та віртуального світів для створення нових середовищ та візуалізацій, де фізичні та цифрові об'єкти співіснують і взаємодіють у реальному часі. MR поділяється на дві основні підмножини:

Доповнена віртуальність (Augmented Virtuality, AV): Технології, які додають елементи реального світу (наприклад, зображення користувача чи його рук) до переважно віртуального середовища. Хоча AV є невід'ємною частиною MR, вона історично мала менший обсяг досліджень та розробок порівняно з AR.

Доповнена реальність (Augmented Reality, AR): Технології, які накладають віртуальну інформацію або графічні елементи на поле зору користувача, не закриваючи його фізичне середовище. AR комбінує реальні та віртуальні об'єкти для покращення та збагачення взаємодії з реальним світом [22]. Завдяки швидкому розвитку мобільних технологій та їхній високій доступності, AR значно перевершила AV у дослідженнях, розробках та поширеності серед широкої громадськості.

3. Розширена реальність (Extended Reality, XR)

Розширена реальність (XR) – це парасольковий термін (umbrella term), що використовується для позначення всіх раніше обговорених форм реальності: VR, AR та MR. Іноді XR також згадується як крос-реальність (Cross-Reality, CR). XR охоплює всі технології, які передбачають

представлення як реальних, так і віртуальних об'єктів у межах одного дисплея або взаємодії. Використання терміна XR є доцільним при обговоренні складних, мультимодальних систем, які інтегрують більше однієї форми реальності (наприклад, система, що функціонує як у VR, так і в AR режимах). Оскільки галузь XR демонструє експоненційне зростання, проектування ефективних та масштабованих інтерфейсів користувача для цих складних, крос-платформних систем вимагає подальшого, цілеспрямованого наукового дослідження.

1.3. Роль швидкого прототипування в проектуванні інтерфейсів розширеної реальності

Прототипування дизайнів є потужним інструментом у дослідницькому процесі, оскільки воно дозволяє валідувати передбачуваний кінцевий вигляд та функціональність системи без необхідності інвестувати значний час та ресурси у повномасштабну розробку. Це забезпечує можливість отримання цінних відгуків на ранніх ітераційних етапах.

1.3.1. Переваги та специфіка інтерактивного прототипування

Сучасне програмне забезпечення, призначене для створення високоякісних дизайнів та швидкого прототипування, часто підтримує імітацію інтерактивності за допомогою каркасів та анімацій (наприклад, імітація взаємодій користувача: перетягування, натискання, наведення). Така імітація створює ілюзію повністю функціонального продукту.

Використання інтерактивного прототипування сприяє більш ітеративному процесу дизайну, дозволяючи прискорене тестування взаємодій користувачів з метою раннього виявлення недоліків дизайну та проблемних моментів (pain points) до початку фази повної технічної реалізації.

Прототипування є особливо критичним при проектуванні інтерфейсів для систем доповненої реальності (AR) та віртуальної реальності (VR), які

використовують шоломи (HMDs). Це зумовлено високою трудомісткістю та ресурсозатратністю технічної реалізації у цих системах, яка часто вимагає використання дорогого та спеціалізованого апаратного забезпечення для тестування. Прототипування дозволяє розробникам візуалізувати перспективу користувача ("through the eyes" view) та просторове розташування інтерфейсу до початку складного етапу впровадження.

Незважаючи на зростання популярності швидкого прототипування у дослідженнях та промисловості (використовуючи таке програмне забезпечення, як Adobe XD, InVision, Framer), програмне забезпечення, спеціалізоване для прототипування додатків XR, залишається обмеженим через новизну та технологічну складність галузі.

Проектування, тестування та створення прототипів додатків XR є складним завданням, оскільки вимагає значних часових та ресурсних витрат на технічне впровадження систем із застосуванням HMD-апаратного забезпечення та просунутих сенсорних технологій. Як наслідок, створення прототипів у AR, VR та змішаній реальності (MR) часто обмежується експертами або особами, які мають просунуті технічні знання та навички програмування.

Хоча дослідники нещодавно опублікували інструменти, спрямовані на розширення поточних методів швидкого прототипування для XR, з метою швидшого моделювання XR-інтерфейсів [11], ці інструменти перебувають на ранніх етапах розвитку. Наразі вони не надають універсальних та стандартизованих методів для збору кількісних даних про взаємодію користувачів, особливо в контексті дистанційного тестування користувачів (online remote user testing).

1.3.2. Проблема дистанційного тестування та вибір методології

Дистанційне тестування прототипів становить додатковий виклик, особливо коли для перегляду та взаємодії потрібне спеціалізоване апаратне забезпечення HMD. Для мінімізації цих обмежень може бути використана

комбінація інструментів для створення каркасів прототипів із спеціалізованим програмним забезпеченням для онлайн-тестування користувачів, що дозволяє збирати дані про їхню взаємодію.

У контексті цієї роботи для створення макетів дизайну використовується Adobe XD. Вибір обумовлений його векторною платформою, яка полегшує перенесення дизайнів на етап технічного впровадження, а також його здатності ефективно використовуватися у сценаріях дистанційного тестування користувачів. Дослідження дистанційних досліджень зручності використання XR із застосуванням прототипів, детально обговорюване в наступних розділах, надає методологічну перевагу в ситуаціях, коли зацікавлені сторони та кінцеві користувачі географічно віддалені від дослідних майданчиків.

1.4. Дослідження існуючих інструментів швидкого прототипування для XR

Оскільки технології XR (Extended Reality) швидко розвиваються, потреба у швидкому та доступному прототипуванні інтерфейсів користувача (UI) для VR, AR та MR стає критичною. Існуючі інструменти для швидкого прототипування XR можна класифікувати за їхнім рівнем занурення та складністю використання. Вони прагнуть подолати традиційні обмеження, пов'язані з високими вимогами до кодування та спеціалізованого обладнання.

1.4.1. Інструменти прототипування на основі скриптингу/ігрових движунів (Expert Level)

Ці інструменти забезпечують максимальну гнучкість, інтерактивність та реалізм, але вимагають глибоких технічних знань.

Unity 3D - провідний крос-платформний ігровий рушій, який є галузевим стандартом для розробки та прототипування майже всіх комерційних XR-додатків. Дозволяє створювати повноцінні 3D-середовища,

інтегрувати складну логіку взаємодії та працювати з усіма основними HMD (Oculus, Vive, HoloLens).



Рис. 1.2. Вигляд ігрового рушія Unity 3D

Хоча Unity 3D є потужним, він вимагає значних навичок програмування (C#) і не є "швидким" у сенсі нетехнічного прототипування.

Unreal Engine (UE) - ще один потужний ігровий рушій, відомий своєю фотореалістичною графікою. Використовує систему візуального скриптингу Blueprints, що дещо спрощує прототипування без написання коду, але все ще вимагає розуміння складної логіки.

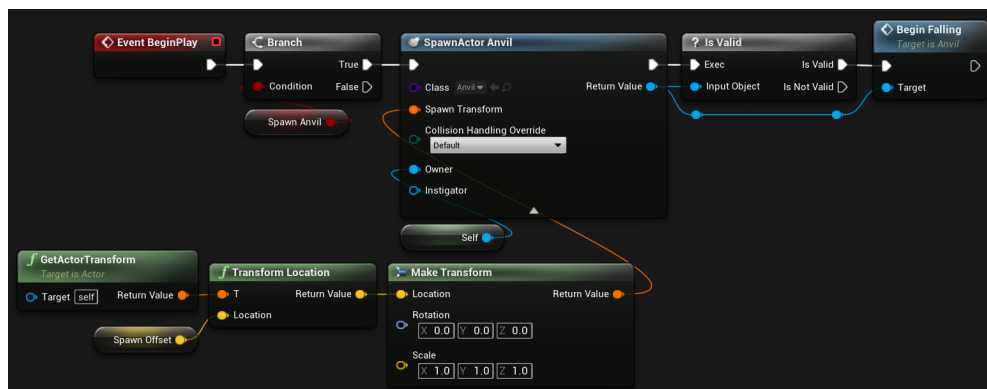


Рис. 1.3. Система візуального скриптингу Blueprints

1.4.2. Інструменти для візуального скриптингу та «No-Code» / «Low-Code» (Intermediate Level)

Ці інструменти безпосередньо спрямовані на прискорення процесу прототипування для дизайнерів, які не є розробниками. Вони частково відповідають новим інструментам, згаданим у [9].

Torch / Reality Composer (для AR) - програми, орієнтовані на мобільну AR, що дозволяють створювати прототипи 3D-сцен і взаємодій без кодування. Дизайнери можуть імпортувати 3D-моделі та встановлювати прості логічні тригери (наприклад, "при натисканні відтворити анімацію").

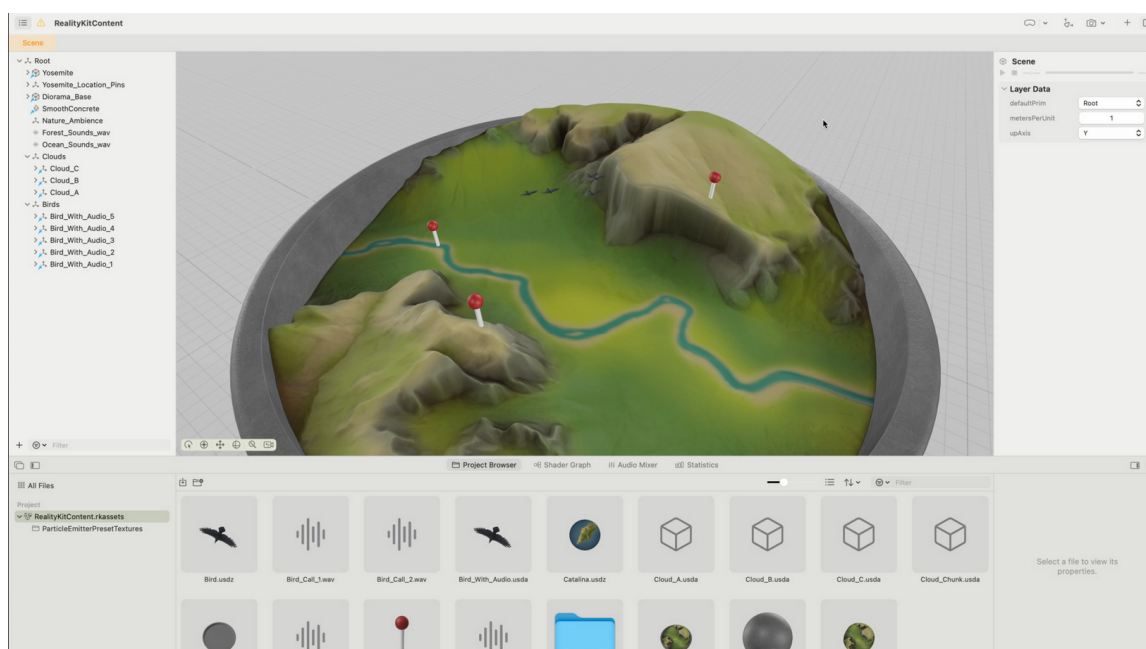


Рис. 1.4. Приклад Reality Composer програми

PrimeSense / Microsoft Maquette - інструменти, розроблені для створення прототипів безпосередньо у VR (in-VR design). Дозволяли дизайнерам малювати та створювати 3D-інтерфейси, використовуючи контролери VR-гарнітури, що є більш інтуїтивним для просторового дизайну. На даний час не активні

Wired Actions (Unity Add-on) - додаток, що дозволяє створювати прототипи всередині Unity за допомогою візуальних вузлів (як Blueprints), мінімізуючи необхідність коду.

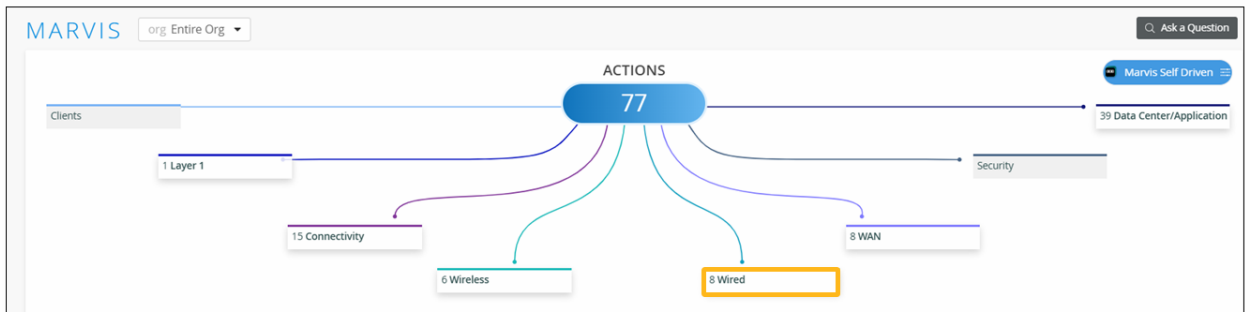


Рис. 1.5. Приклад роботи Wired Actions

A-Frame / Babylon.js - фреймворки, що дозволяють створювати VR/AR-досвід для вебу (WebXR) за допомогою простого HTML-подібного синтаксису. Перевагою є висока доступність, не вимагає спеціального програмного забезпечення для перегляду (доступно через браузер).

1.4.3. Інструменти на Основі 2D-Макетів та Імітації (Conceptual Level)

Ці інструменти є традиційними, але можуть використовуватися для концептуального дизайну XR-інтерфейсів, особливо для навігації та плоских 2D-елементів (що відповідає вашому використанню Adobe XD).

Adobe XD, Figma, Sketch використовуються для XR з метою створення "плоских" 2D-елементів UI (меню, кнопки, панелі, іконки), які пізніше будуть імпортовані та розміщені у 3D-просторі.

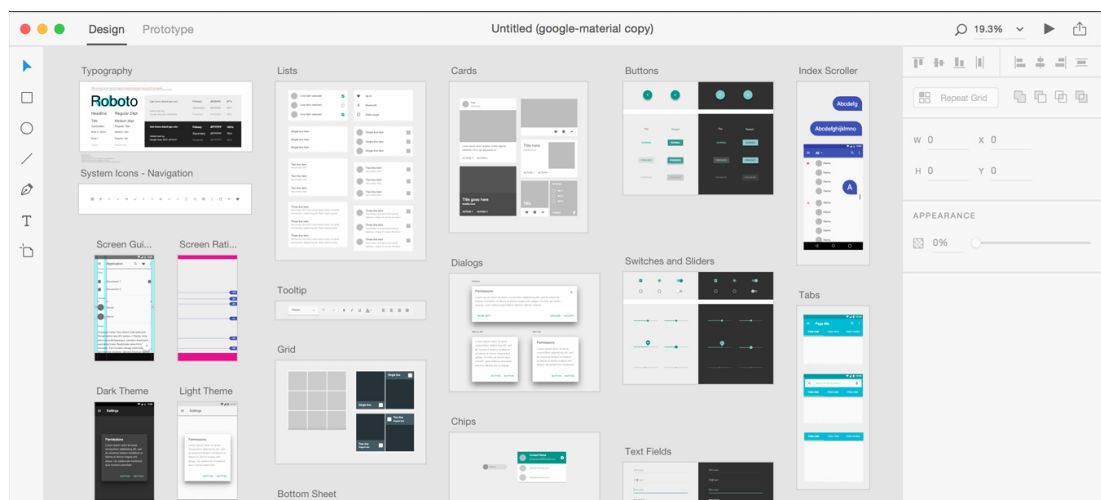


Рис. 1.6. Інструмент Adobe XD

Перевагами є швидкість, висока якість векторної графіки, легкість у зборі відгуків, особливо в дистанційному тестуванні.

1.4.4. Нові інструменти для швидкого моделювання XR

Дослідження активно розробляють нові, більш спеціалізовані інструменти, які заповнюють прогалину між 2D-макетами та повною розробкою. Основні наведені в таблиці 1.2.

Таблиця 1.2.

Опис нових інструментів для швидкого моделювання XR

Фокус дослідження	Мета інструменту	Ключова інновація	Проблематика
"Mid-Fidelity XR Prototyping"	Швидке моделювання 3D-інтерфейсів та просторових взаємодій.	Дозволяють дизайнерам працювати з 3D-простором та маніпулювати елементами без Unity/Unreal.	Обмежена інтерактивність та складність логіки.
Крос-платформне прототипування	Створення єдиного прототипу, що адаптується до VR, AR та настільних комп'ютерів.	Використання абстрактних компонентів UI, які рендеряться по-різному залежно від режиму реальності.	Складність збору кількісних даних про поведінку користувачів та просторову метрику.
Тестування на основі НМД	Доступність для нетехнічних користувачів для перегляду та тестування прототипів у НМД.	Експорт прототипу безпосередньо у формат, який може бути запущений на НМД одним клацанням.	Відсутність стандартизації для дистанційного збору даних.

Незважаючи на існування потужних рушіїв (Unity, Unreal) та простих 2D-інструментів (Adobe XD), проблема швидкого, доступного та науково обґрунтованого прототипування XR (особливо для крос-платформних систем із збором кількісних даних) залишається актуальною науковою прогалиною. Дане дослідження, зосереджене на використанні Adobe XD для створення 2D-елементів та подальшого аналізу їхньої масштабованості та зручності

використання на крос-платформній системі, що є методологічно виправданим у цьому контексті обмежень.

1.5. Концептуалізація та дослідницькі прогалини масштабованих крос-реальних систем

Дослідження, що стосуються масштабованості систем розширеної реальності (XR), залишаються мінімальними, що підкреслює існуючу наукову прогалину. Проте, у роботі [12] вводиться концепція масштабованої розширеної реальності (Scalable Extended Reality, XR*).

Автори визначають XR* як концепцію, що передбачає масштабування як між різними дисплеями, так і між різними ступенями віртуальності. Основна мета цієї масштабованості полягає у підвищенні запам'ятовуваності (memorability), що дозволяє користувачам легше перемикатися між різними підкатегоріями XR та пристроями без необхідності значного психічного навантаження та повторного навчання.

Стверджують, що відсутність масштабованості є обмежуючим фактором для широкого впровадження технології XR, попри її значний потенціал у таких галузях, як професійне навчання та робочі середовища. Вони ідентифікують кілька тем для майбутніх досліджень, необхідних для підтримки масштабованих XR-систем. Зокрема, вони наголошують на необхідності розробки технік взаємодії, які залишаються інтуїтивними для користувачів навіть при перемиканні між різними пристроями або ступенями віртуальності. Автори пропонують консолідувати існуючі незалежні галузі досліджень для повноцінного створення та вивчення масштабованих систем XR.

Аналогічно, в роботі [13] пропонують створення простору дизайну, орієнтованого на такі мультимодальні додатки, але визначають його як Крос-Реальність (Cross Reality, CR). Вони визначають CR як використання

інструментів для вирішення проблем, які розташовані у різних точках шкали реальність-віртуальність Мілграма.

Дослідники наголошують на необхідності вивчення одноосібних додатків CR та окреслюють потенційні підгалузі для майбутніх досліджень у цій сфері. Вони також підкреслюють труднощі, з якими стикаються розробники та дослідники через відсутність структурованих настанов при прийнятті рішень щодо компонентів додатків та інтерфейсів користувача в AR, VR та інших режимах.

Хоча наукові роботи підтверджують потребу в дослідженнях та розробці додатків, що функціонують у кількох реальностях, критично мало або зовсім немає наукових праць, які б надавали практичні рекомендації щодо дизайну інтерфейсів користувача (UI), які успішно масштабуються та працюють у різних XR-режимах.

Поточні інтерфейси для багатомодальних систем часто проектуються на основі існуючих обмежень апаратного та програмного забезпечення, зручності (expendiency) або швидкості розробки, і, як правило, пріоритет дизайну UI залишається низьким. У міру того, як кількість додатків, які використовують більше одного типу реальності, продовжує зростати, додаткові дослідження, що стосуються цих просторів дизайну та, зокрема, інтерфейсів, розроблених для цих систем, є нагально необхідними.

1.6. Розробка принципів дизайну інтерфейсів користувача для розширеної реальності

Існуючий науковий доробок включає створення основоположних принципів дизайну інтерфейсів користувача (UI), найвідомішими серед яких є десять евристик зручності використання Нільсена [19]. Проте ці евристики мають обмежену застосовність, оскільки вони не були розроблені з урахуванням специфіки інтерфейсів та технологій розширеної реальності (XR).

1.6.1. Специфічні виклики UI в XR-системах

На відміну від традиційних двовимірних непросторових інтерфейсів, XR-інтерфейси, зокрема ті, що використовують шоломи з дисплеями (HMD), мають принципові відмінності у методах перегляду та взаємодії користувачів з інформацією. Це породжує унікальний набір дослідницьких проблем.

Ключові особливості XR-систем, які ускладнюють застосування традиційних евристик, включають:

- Розширені методи введення, відстеження руху та напрямку погляду.
- Просторова Взаємодія: 360-градусний перегляд інформації та маніпулювання тривимірними об'єктами.
- Мультисенсорна інтеграція - дослідження продовжуються щодо комбінації гаптичних можливостей, просторового звуку та додаткових сенсорів у поєднанні з XR-платформами.

Відтак, розробка рекомендацій з дизайну, специфічних для різних підмножин XR, є важливим внеском у наукове співтовариство, спрямованим на розширення та адаптацію поточних, загальноновизнаних принципів дизайну. Такі керівні принципи можуть бути використані для акцентування на анти-патернах (чого слід уникати) при проектуванні UI, допомагаючи мінімізувати відомі проблеми, що виникають у різних XR-системах.

1.6.2. Існуючі спеціалізовані рекомендації

Деякі дослідницькі зусилля вже були спрямовані на створення рекомендацій, специфічних для окремих категорій XR:

- AR-специфічні рекомендації призначені для практиків AR, з метою ефективнішого вирішення проблем проектування, які не охоплюються евристичними Нільсена [19]. Ці евристики були розроблені із застосуванням методу афініті-діаграмування на основі існуючих на той час рекомендацій у сфері AR.

- Рекомендації для HMD це одинадцять рекомендацій, спеціально орієнтованих на апаратне забезпечення HMD. Це дослідження синтезувало

інформацію з понад 60 різних ресурсів, включаючи опубліковані академічні роботи та онлайн-документацію від провідних розробників HMD.

Незважаючи на ці зусилля, дослідники виявили, що існуюча робота обмежена у своїй здатності бути практично застосованою, часто через недостатність прикладів застосування концепцій у розроблених XR-додатках [21]. Ця прогалина часто призводить до запозичення з інших доменів, наприклад, до застосування загальних рекомендацій HCI до AR, як це було зроблено в [23].

Для вирішення ідентифікованих обмежень, ця робота має на меті надати рекомендації, які можуть бути застосовані саме на початкових етапах проектування для масштабування дизайнів інтерфейсів користувача між модальностями XR.

Хоча проведено деякі дослідження користувацького досвіду та рекомендацій щодо дизайну UI в окремих середовищах AR, VR та MR HMD, додатки розширеної (XR) та крос-реальності (CR), які дозволяють користувачам використовувати ті самі інструменти в різних реальностях та типах HMD, не забезпечені існуючими керівними принципами та евристичними. Загалом, дослідження, що стосуються створення та практичного застосування рекомендацій для розробки масштабованих багатомодальних інтерфейсів, залишаються на ранніх етапах розвитку.

Висновки до розділу

У першому розділі було сформовано цілісне бачення предметної області масштабування інтерфейсів у віртуальній та доповненій реальності, що дозволило визначити фундаментальні властивості XR-систем і особливості їхнього інтерфейсного забезпечення. У ході дослідження з'ясовано, що крос-платформенність є ключовим чинником ефективної масштабованості, оскільки різні XR-пристрої мають неоднакові обчислювальні можливості, сенсорні засоби та способи подання інформації.

Аналіз континууму віртуальності дав змогу розмежувати технології від доповненої до повністю віртуальної реальності, що, у свою чергу, визначає специфіку дизайну UI і підходи до організації взаємодії. Особлива увага була приділена швидкому прототипуванню, яке в XR-проєктах виступає інструментом зниження витрат, прискорення ітераційного циклу та підвищення точності користувацьких рішень. Проведений огляд інструментів прототипування показав, що галузь характеризується значною фрагментованістю: окремі засоби орієнтовано на скриптинг, інші — на візуальне моделювання або концептуальні макети. Така різноманітність створює як широкі можливості, так і певні труднощі щодо вибору універсального підходу в умовах масштабованих проєктів. Також у розділі було виявлено концептуальні прогалини, пов'язані з відсутністю уніфікованих методів моделювання UI/UX для крос-реальних систем, що перешкоджає узгодженості мультимодальних компонентів.

РОЗДІЛ 2. МЕТОДОЛОГІЯ ПРОЕКТУВАННЯ ІНТЕРФЕЙСУ ЗА ДОПОМОГОЮ ШВИДКОГО ПРОТОТИПУВАННЯ ДЛЯ ВІРТУАЛЬНОЇ РЕАЛЬНОСТІ

2.1. Концептуалізація персонажів користувачів та афініті діаграмування

Цей розділ деталізує методологію, використану для створення макетів інтерфейсу користувача (UI) із застосуванням програмного забезпечення швидкого прототипування Adobe XD. Прототипи були розроблені на основі емпіричних даних у сфері користувацького досвіду (UX). У цьому розділі представлені персонажі користувачів та функціональні вимоги, які слугували початковою основою для дизайну, описано процес проектування, включно із застосуванням афініті-діаграмування, та наведено опис створених UI-макетів у Adobe XD.

Персонажі користувачів є "фіктивними репрезентаціями користувачів, створеними для втілення поведінки та мотивацій, які можуть виявляти реальні користувачі". Персонажі є важливим інструментом для дизайнерів UX, оскільки вони дозволяють чітко визначити критичні функції та врахування макету вже на етапі прототипування. Крім того, використання персонажів надає значні переваги процесу прототипування, безпосередньо інформуючи дизайнерські рішення та забезпечуючи актуальність та релевантність створюваних дизайнів для цільової аудиторії.

Перед початком розробки макетів інтерфейсу користувача було сформульовано первинний набір персонажів та вимог, щоб забезпечити початкову спрямованість проектування, яка б відповідала потребам кінцевих користувачів, які використовують інструмент візуалізації та взаємодії.

Через таку значну варіативність технічної обізнаності, система вимагає інтерфейсів, які функціонують ефективно для широкого діапазону

користувачів. Для вирішення цієї проблеми були розроблені два контрастні персонажі:

- Один персонаж із високорозвиненими технічними знаннями.
- Інший персонаж із мінімальними технічними знаннями.

Такий підхід забезпечив охоплення різноманітності експертизи серед кінцевих користувачів. Ці початкові персонажі, вимоги та припущення, які детально описані в наступних підрозділах, будуть згодом емпірично верифіковані.

Афіниті-діаграмування (affinity diagramming), або метод спорідненості, є якісною дослідницькою технікою, яка широко використовується в проєктуванні, орієнтованому на користувача (User-Centered Design, UCD), та аналізі даних інтерв'ю. Метод дозволяє організувати велику кількість неструктурованих даних (ідей, спостережень, цитат користувачів) у логічні групи на основі їхньої природної спорідненості або зв'язку.

Афіниті-діаграмування слугує мостом між необробленими, хаотичними даними, зібраними під час якісних досліджень (таких як інтерв'ю з експертами UX), та чіткими, дієвими дизайнерськими вимогами та інсайтами.

Ключові етапи методу

1. Збір даних. Запис усіх необроблених даних (цитат, спостережень, ідей) на окремих картках або цифрових нотатках.
2. Групування (Clustering). Фізичне або віртуальне переміщення нотаток, які стосуються однієї теми чи проблеми, разом.
3. Створення заголовків (Labeling). Надання кожній групі узагальнюючої тематичної назви (наприклад, "Проблеми кінематозу", "Потреби в масштабованості UI", "Вимоги до голосового введення").
4. Синтез інсайтів. Перетворення цих тематичних груп у конкретні дизайнерські принципи, евристики або функціональні вимоги.

У контексті даної роботи, де досліджуються крос-платформні XR-системи, афіниті-діаграмування має вирішальне значення.

Особливості афініті-діаграмування в даному дослідженні

Фокус	Роль афініті-діаграмування
Мультимодальність	Допомагає систематизувати конфліктні вимоги (наприклад, "потрібен фізичний контролер у VR, але потрібні жести у AR") та звести їх до єдиних масштабованих принципів.
Просторова взаємодія	Групує спостереження щодо ергономіки та комфорту (наприклад, "меню занадто далеко", "важко взаємодіяти жестом") для визначення оптимальних просторових зон для розміщення UI-елементів.
Створення персонажів	Дозволяє об'єднати поведінкові патерни (наприклад, "швидкість прийняття рішень офіцером", "потреба у навчанні рядового складу") у дві або три чіткі категорії користувачів, які ви використовуєте для валідації дизайну.
Екстракція евристик	Дозволяє виявити повторювані проблемні моменти (наприклад, VIMS, дезорієнтація) та сформулювати на їхній основі нові, специфічні для XR евристики або рекомендації, які доповнюють традиційні принципи Нільсена.

Таким чином, афініті-діаграмування є систематичним інструментом аналізу якісних даних, що забезпечує прямий, обґрунтований перехід від необроблених думок експертів до структурованих технічних і дизайнерських специфікацій для прототипів UI в Adobe XD.

2.2. Основні етапи проектування початкових ескізів інтерфейсу та віртуального дизайну

2.2.1. Проектування ескізів

Дизайни прототипів інтерфейсу користувача були реалізовані за допомогою спеціалізованого програмного забезпечення для швидкого прототипування (Adobe XD). Вибір цього інструменту забезпечив можливість створення деталізованих каркасів потоків завдань та подальшої їхньої інтеграції з програмним забезпеченням для тестування користувачів з

метою проведення дистанційного тестування зручності використання (remote usability testing).

Процес створення макетів охоплював такі етапи:

- 1) початкове концептуальне проектування з використанням ручки та паперу;
- 2) прийняття рішень щодо візуального дизайну, включаючи колірну палітру та іконографіку;
- 3) визначення ключових завдань для каркасування;
- 4) перенесення та реалізація в програмному забезпеченні Adobe XD для детальної візуалізації передбачуваного кінцевого продукту.

Початкові ескізи, що відображали логіку потоку завдань, були створені вручну (ручка та папір). Хоча на цьому попередньому етапі головним дослідником визначалося загальне розташування меню, результати та рішення щодо цих макетів підлягали валідації експертами з предметної області (СПО). Попередні ескізи дизайну представлені на рисунку 2.1.

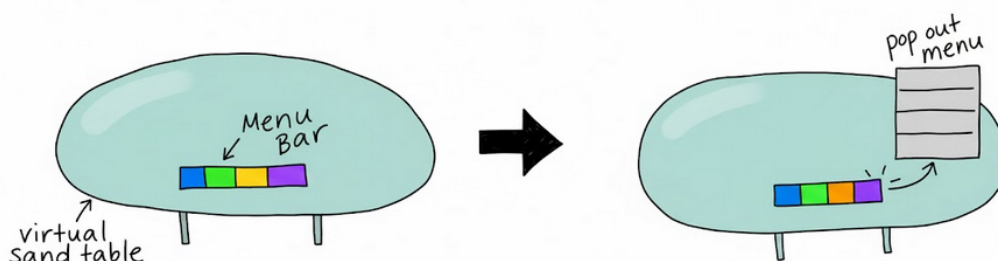


Рис. 2.1. Початкові ескізи прототипів які демонструють відкриття додаткових меню після вибору елемента.

Візуальний дизайн UI ґрунтувався на принципах Material Design від Google. Структура та макет головної панелі меню базувалися на шаблоні навігаційної панелі (Navigation Rail), тоді як окремі внутрішні панелі меню були змодельовані за елементами навігаційної панелі, списку та перемикача.

Іконографія була обрана з бібліотеки Material Icons та Symbols від Google Fonts. Іконки були максимально узгоджені з їхніми відповідними

завданнями або заголовками меню, з мінімальними модифікаціями, коли виникала потреба передати більш унікальне семантичне значення.

Кольорова палітра була обрана з темною основою та високою контрастністю для досягнення двох цілей:

- 1) мінімізувати візуальне відволікання від основного простору пісочного столу;
- 2) підвищити читабельність елементів інтерфейсу в умовах доповненої реальності.

2.2.2. Визначення завдань для прототипування макетів

Набір тестових завдань був сформульований на основі припущень щодо вимог до планування, які були визначені для обох створених персонажів користувачів. Вимога до редагування була декомпозирована на два окремі завдання, оскільки огляд тактичних планів підтверджує необхідність роботи з різними графічними зображеннями та символами. Отже, додавання анотаційної інформації було розділено на окремі завдання для тактичних символів та графічних елементів.

Таблиця 2.2.

Перелік завдань для символів та графічних елементів

№	Завдання	Опис дії
Завдання 1	Додати "Символ"	Знайти та створити новий Символ. Використовувати налаштування за замовчуванням без введення назви.
Завдання 2	Додати "Графіку" (A→B)	Знайти та створити нову Графіку. Використовувати налаштування за замовчуванням (стрілку), вказавши послідовно точку A та точку B для її розміщення.
Завдання 3	Перемістити Символ	Вибрати та перемістити існуючий Символ з позначки A до позначки B за допомогою механізму перетягування.
Завдання 4	Увімкнути перегляд	Знайти та активувати видимість моделей транспортних засобів.
Завдання 5	Зберегти Сценарій	Знайти та вибрати опцію "Зберегти сценарій", підтвердивши дію.

П'ять обраних завдань демонструють варіативність за кількістю кроків, необхідних для їхнього завершення. Ця різниця зумовлена розташуванням елементів в ієрархії меню, необхідністю введення координат для розміщення складної графіки, а також включенням таких функцій, як перетягування (drag-and-drop), які не вимагають навігації по системі меню. Таким чином, вибрані завдання вимагають від користувача навігації та взаємодії з UI різними способами. Наприклад, завдання 2 вимагає переходу до відповідного елемента меню, а потім вибору двох контрольних точок для визначення форми та розміру графіки. Натомість завдання 3 моделює зміну розташування існуючого тактичного символу, що реалізується за допомогою рухів перетягування.

2.2.3. Прототипування засобами Adobe XD

На основі ієрархії системи меню, описаної в на початку цього розділу та візуальних елементів дизайну, компоненти Material Design були імпортовані в Adobe XD та використані для побудови інтерфейсу користувача. Material Design ґрунтується на метафорі фізичного світу, використовуючи принципи реалістичних матеріалів, світла та руху для створення візуальної ієрархії та зрозумілої взаємодії.

Спочатку була реалізована головна панель меню у вигляді набору з чотирьох перемикачів у навігаційній панелі. Потім кожна внутрішня панель меню та відповідні підпанелі були розроблені як навігаційна панель зі списком елементів. На цьому етапі були встановлені кольори UI-елементів (колір тексту, фоновий колір).

Після створення меню вони були інтегровані в тривимірний ландшафт. Фонове зображення для цього було отримано шляхом запуску існуючого програмного забезпечення в режимі віртуальної реальності, імпортування тривимірної місцевості та захоплення знімка екрана порожнього фону та столу планування місії. Знімки екрана дизайнерських прототипів у Adobe XD представлені на рисунку 2.1.

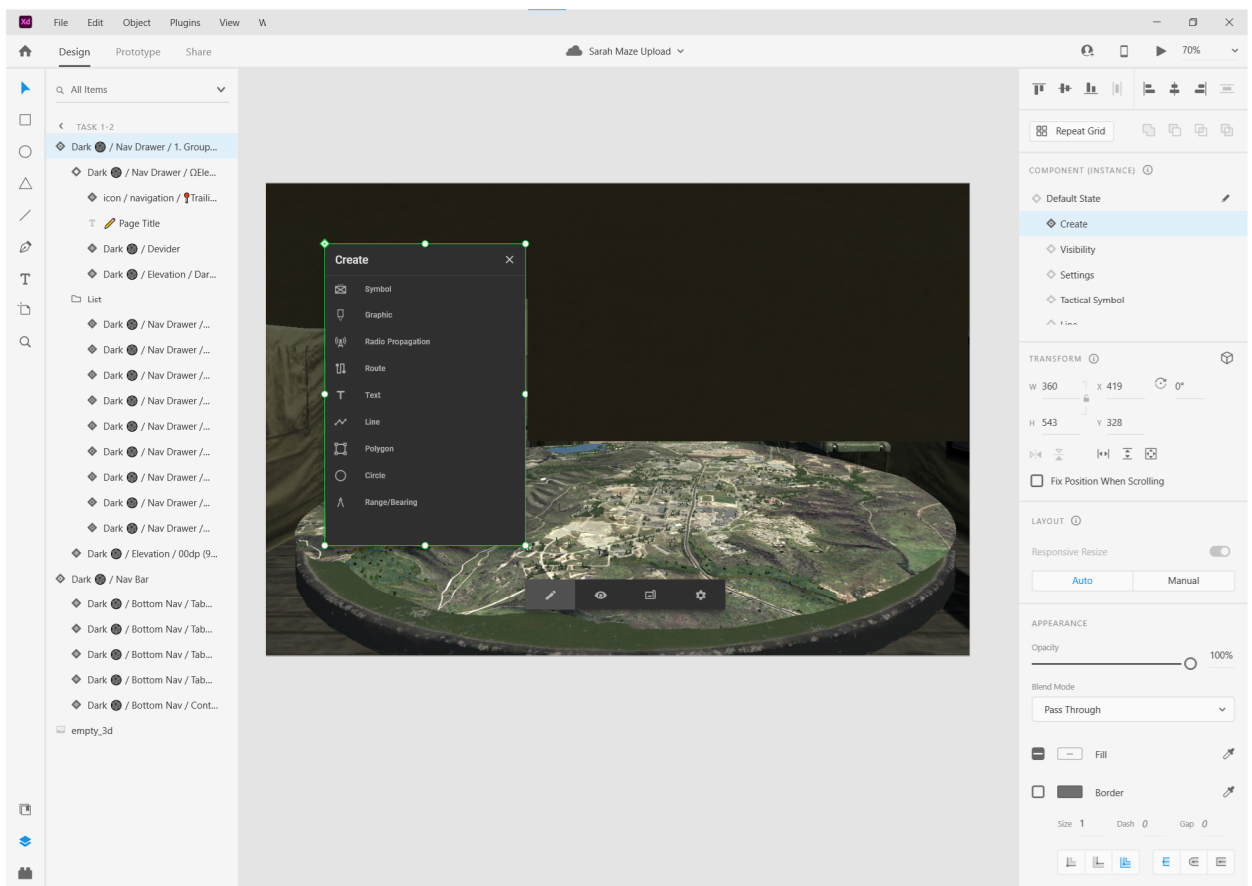


Рис. 2.1. Інтерфейс програмного забезпечення Adobe XD, який забезпечує детальне проектування макетів інтерфейсу користувача

Для створення інтерактивних та реалістичних прототипів були використані функції автоанімації (Auto-Animate) Adobe XD для різних типів введення, включаючи клацання, наведення та перетягування. Були додані стани наведення для елементів UI та анімації перетягування елементів. На рис. 2.2. показано приклад взаємодії, де клацання по іконці меню "Редагувати" ініціює відкриття підменю з кількома опціями редагування та анотації.

Потік взаємодії для всіх п'яти завдань візуалізовано на рисунку 2.3, який відображає необхідний лінійний шлях для завершення кожного завдання. Деталізований потік руху, що ілюструє переходи між початковим та кінцевим екранами для завдання 2, представлений на рисунку 2.4. Цей потік передбачає, що користувач послідовно відкриє меню 'Редагувати',

вибере 'Графіку', 'Створити', а потім вкаже дві точки на столі для розміщення графічного елемента (стрілки) за замовчуванням.

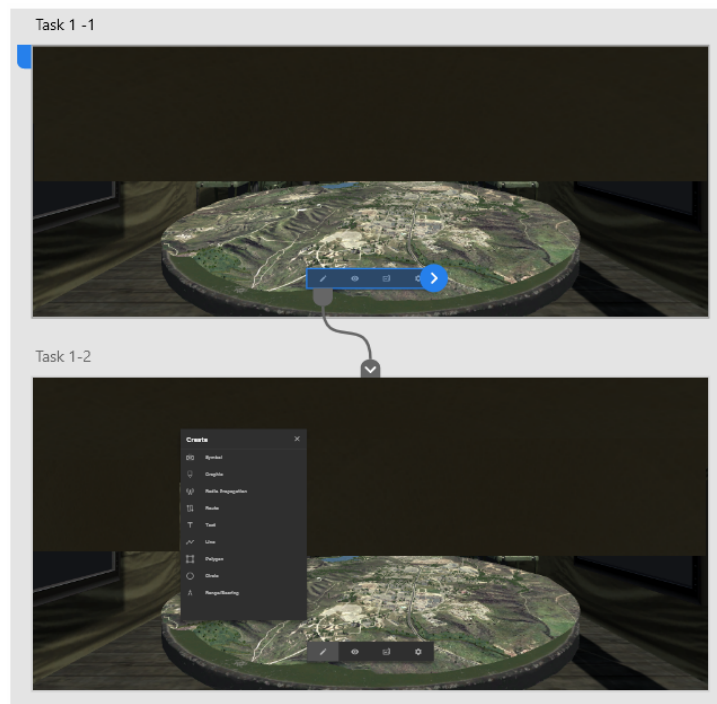


Рис. 2.2. Знімок екрана Adobe XD, що демонструє додану інтерактивність макета для відкриття меню редагування



Рис. 2.3. Потіки інтерактивності Adobe XD для п'яти завдань планування, від їхніх початкових екранів до завершення завдання

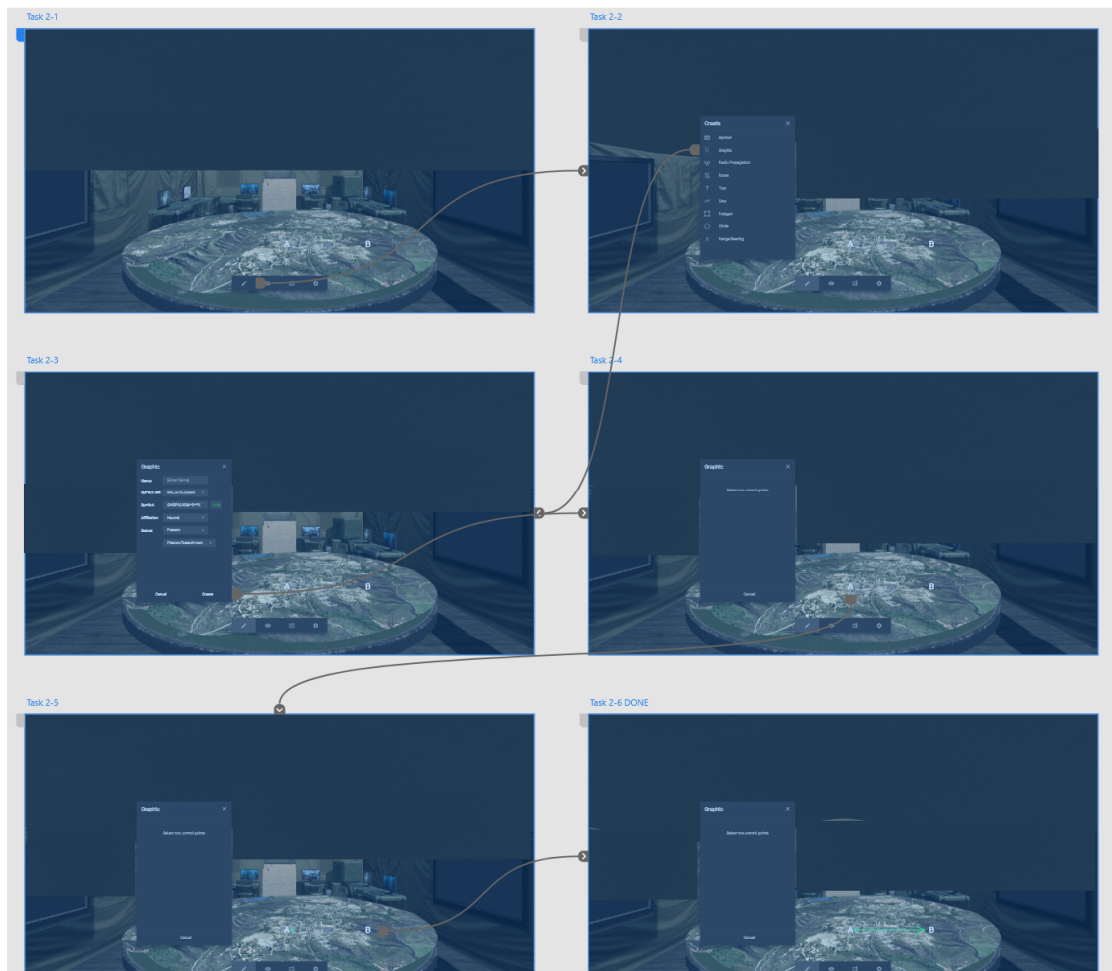


Рис. 2.4. Потік інтерактивності Adobe XD для завдання 2, від початкового екрана до завершення

Для забезпечення дистанційного збору даних було використано платформу швидкого тестування Maze. Maze — це веб-платформа, що дозволяє створювати налаштовані користувацькі сценарії тестування. Maze — це веб-орієнтована платформа для дистанційного, немодерованого тестування користувачів (remote, unmoderated user testing), яка спеціалізується на швидкій валідації прототипів інтерфейсу користувача (UI) та потоків взаємодії. Вона слугує ключовим інструментом для дослідників та дизайнерів, які прагнуть збирати кількісні та якісні дані про зручність використання (usability) на ранніх етапах розробки. Maze підтримує пряму інтеграцію через плагіни з популярними дизайнерськими програмами, такими як Adobe XD, Figma та Sketch

Завдяки спеціалізованому плагіну Maze для Adobe XD завершений інтерактивний прототип та відповідні потоки завдань були імпортовані в Maze. Це дозволило здійснити дистанційний збір даних щодо зручності використання.

Платформа Maze збирає різноманітні метрики, що робить її ефективною для кількісної оцінки UI (таблиця 2.3).

Таблиця 2.3.

Метрики Maze

Метрика Maze	Опис та значення
Коефіцієнт успіху (Success Rate)	Відсоток користувачів, які успішно завершили завдання, дотримуючись очікуваного шляху.
Прямота (Directness)	Метрика, що вимірює, наскільки близько фактичний шлях користувача відповідав очікуваному шляху (тобто, чи були зайві кліки або обхідні шляхи).
Середній час (Time on Task)	Час, витрачений користувачами на виконання завдання.
Теплові карти (Heatmaps)	Візуальне відображення місць, де користувачі клікали. Це допомагає виявити "мертві зони" (місця, де очікували кнопку, але її немає) або елементи, що відволікають.
Записи сесій (Recordings)	Запис екрана користувача під час проходження тесту, що допомагає якісно зрозуміти його поведінку.

Використання Maze для тестування прототипів, створених у Adobe XD, є методологічно виправданим, оскільки:

1) дозволяє дистанційно протестувати велику вибірку користувачів, що є критично важливим, якщо зацікавлені сторони (наприклад, військові експерти) географічно розподілені.

2) надає чіткі кількісні показники (успіх, час, прямота) для об'єктивної оцінки зручності використання, що є необхідним для наукового дослідження.

3) дозволяє швидко ідентифікувати проблеми в потоці завдань, перш ніж інвестувати час у дорогу та ресурсозатратну технічну реалізацію у системі XR.

На рисунку 2.5 показано налаштування завдань 1-5 у програмному забезпеченні Maze, включаючи додавання інструкцій та визначення очікуваних шляхів користувачів.

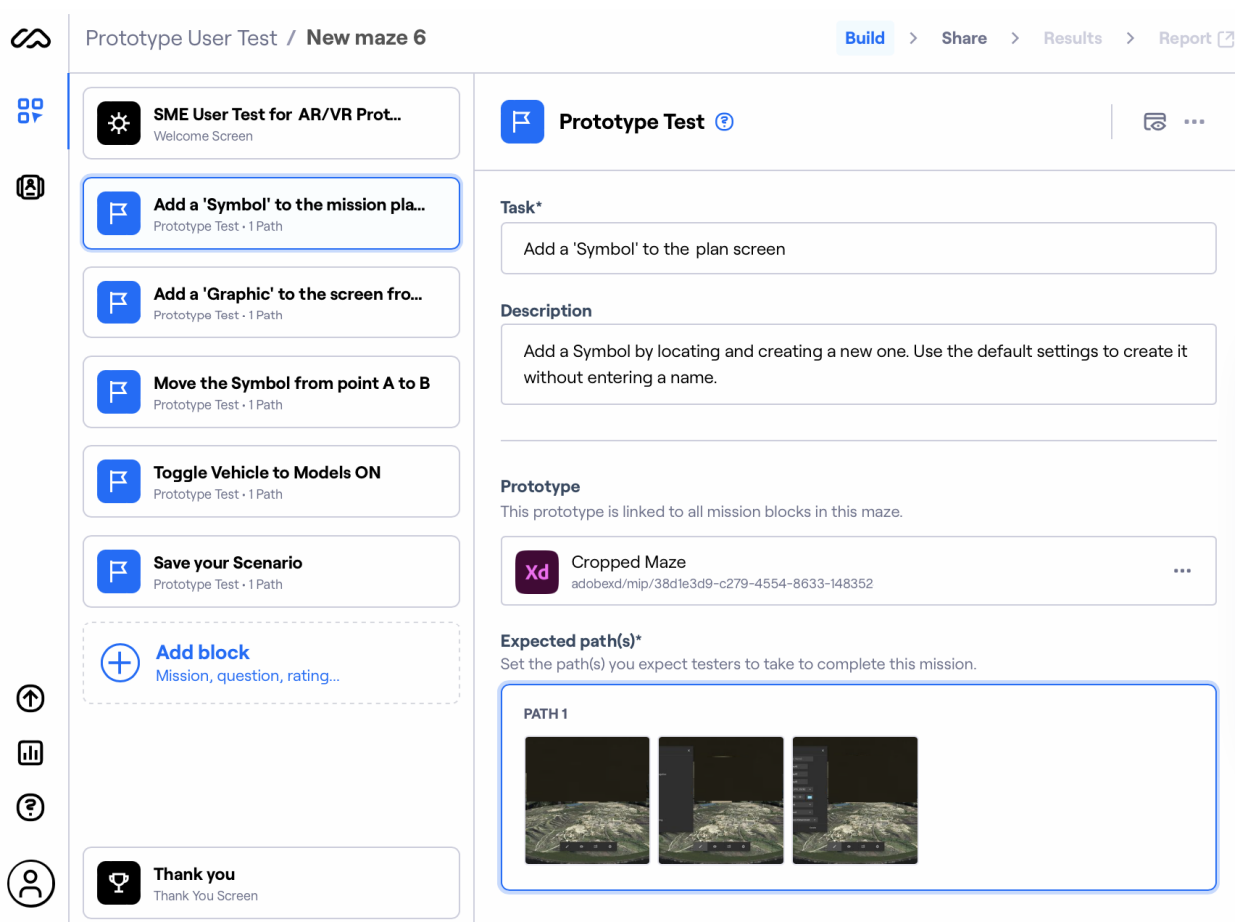


Рис. 2.5. Знімок екрана Maze, що демонструє додавання описів завдань та очікуваних шляхів

Отже, в даному підрозділі описано методологічний процес проектування, використаний для створення початкових каркасних прототипів. Етапи включали визначення персонажів користувачів та функціональних вимог для початкової спрямованості дизайну. Афініті-діаграмування застосовувалося як інструмент для кластеризації та організації схожих елементів у групи, які пізніше трансформувалися у макети меню. На основі вимог до планування місії було визначено п'ять тестових завдань для проектування та оцінки. Ці макети UI були створені в Adobe XD та

інтегровані з платформою Maze для забезпечення дистанційного збору даних щодо зручності використання дизайнів. Результати початкового тестування зручності використання, проведеного з використанням цього прототипу, будуть детально розглянуті в наступному розділі.

2.3. Методологія проектування інтелектуальних віртуальних агентів у середовищах змішаної реальності

Змішана реальність (MR - Mixed Reality) являє собою технологію, яка забезпечує синтез реального та віртуального світів, створюючи нові гібридні середовища, де користувачі можуть здійснювати взаємодію в реальному часі. Здатність накладати комп'ютерну графіку на реальний світ із безшовною взаємодією між віртуальними та фізичними об'єктами робить технологію MR придатною для широкого спектру застосувань.

Сучасні комерційні MR-пристрої, такі як Microsoft HoloLens, Magic Leap та Meta 2, пропонують новітні механізми взаємодії користувачів, які використовують жести рук, голосові команди та відстеження погляду [11].

Протягом десятиліть системи на основі агентів привертають значну увагу дослідників завдяки їхньому потенціалу у різноманітних прикладних сферах [14, 15]. Дослідження [12] представило новий тип антропоморфного агента, функціональність якого реалізована у 3D-просторі, де відбувається безшовне злиття реального та віртуального середовищ. У роботі [13] створено багатомодального віртуального гуманоїдного асистента, призначеного для імітації діалогів віч-на-віч із користувачами у віртуальному середовищі.

Ідеальний агент має точно імітувати людську поведінку. Дослідження [6] присвячене моделюванню людської поведінки для віртуальних агентів. Комплексний огляд сучасного стану використання віртуальних людей у застосунках культурної спадщини представлено у [15], а їхню корисність для

поширення нематеріальної культурної спадщини у віртуальних середовищах підкреслено у [16].

Сучасні дослідження зосереджені на створенні правдоподібного персонажа чи агента, який частково демонструє емоційні та поведінкові здібності під час взаємодії з навколишнім середовищем або користувачами. Проте, існують припущення, що цього рівня реалізму недостатньо для значного покращення інтерактивного досвіду користувача.

Детальне сучасне дослідження архітектури на основі агентів та проблем інтелектуальних віртуальних агентів представлено у [17]. Однак більшість існуючих рішень орієнтовані на стаціонарні системи на базі ПК та часто страждають від повторюваної доставки контенту.

Поточне покоління носінних MR-пристроїв забезпечує комфортніший та більш захоплюючий досвід. Цей досвід може бути додатково оптимізований шляхом інтеграції художньо створеної 3D віртуальної людини, інтелектуального чат-бота, системи розпізнавання об'єктів та інтуїтивної техніки взаємодії.

Відповідно, метою поточного дослідження є представлення методології багатомодальної взаємодії більше ніж одного агента у MR-середовищі, з особливим фокусом на застосуваннях для музейних експозицій та ботанічних садів.

Пропонована система включає такі інтегровані компоненти:

- Віртуальний персонаж.
- Багатомодальна взаємодія.
- Предметно-орієнтований чат-бот.
- Система розпізнавання об'єктів.

Основна мета цієї інтегрованої системи полягає у забезпеченні безшовної, інтерактивної, захоплюючої та не повторюваної взаємодії. Основними перевагами пропонованої архітектури є її гнучкість та адаптивність, досягнуті завдяки модульному підходу до розробки. Ця система надає користувачеві можливість сприймати реальний світ,

доповнений цифровою інформацією, яка передається через дружнього віртуального персонажа, що створює відчуття наявності особистого екскурсовода для збагачення досвіду.

Висновки до розділу

Другий розділ був присвячений побудові методології проектування інтерфейсу XR-систем за допомогою швидкого прототипування, що стало ключовим елементом розробки структурованого ітераційного процесу. На основі концептуалізації персонажів користувачів вдалося створити узагальнені моделі потреб, поведінкових сценаріїв та очікувань різних груп користувачів, що значно підвищило точність подальших рішень у сфері UI/UX. Застосування афінити-діаграмування сприяло впорядкуванню отриманих даних і виявленню тематичних областей, які визначають логіку структурування інтерфейсу. Розроблення ескізів стало не лише першим візуальним кроком у проектуванні, а й способом швидкої перевірки концептуальних припущень щодо зручності та послідовності взаємодії. Надалі відбувалося проектування макетів, спрямоване на формування сценаріїв прототипування, які дозволили протестувати користувацькі потоки та визначити точки навантаження, можливі помилки і місця для оптимізації. Використання Adobe XD надало можливість створити інтерактивні прототипи, які зберігають структуру XR-інтерфейсу та забезпечують максимальне наближення до реальної поведінки системи. Особливо важливим елементом розділу стала розробка методології проектування інтелектуальних віртуальних агентів, що поєднує підходи до моделювання поведінки, побудови діалогових структур і використання мультимодальних каналів взаємодії. Ця методологія дозволила інтегрувати когнітивні принципи в інтерфейс і сформувавати основу для створення природної взаємодії між користувачем і системою. Загалом, у розділі було сформовано комплексний процес від первинного аналізу потреб до створення готових

інтерактивних прототипів, що дозволяє ефективно масштабувати XR-інтерфейси в залежності від вимог конкретного проєкту. Отримані результати забезпечили практичну основу для реалізації архітектурних рішень, поданих у наступному розділі.

РОЗДІЛ 3. РЕАЛІЗАЦІЯ МОДЕЛЕЙ ТА МЕТОДІВ МАСШТАБУВАННЯ МУЛЬТИМОДАЛЬНИХ ІНТЕРФЕЙСІВ У ВІРТУАЛЬНІЙ ТА ДОПОВНЕНІЙ РЕАЛЬНОСТІ

3.1. Імплементация дизайну інтерфейсу в мультимодальну архітектуру візуалізації

Цей розділ представляє етапи реалізації дизайнів інтерфейсу користувача (UI), розроблених на етапі прототипування, в архітектуру просторової візуалізації та взаємодії в модальностях віртуальної реальності (VR) та доповненої реальності (AR). Процес розробки охоплював глибинний аналіз існуючої кодової бази, адаптацію та фіналізацію макету інтерфейсу користувача з урахуванням обмежень цільової платформи, а також опис методів збору даних протоколу, які були імплементовані для емпіричної оцінки.

3.1.1. Аналіз та декомпозиція існуючої архітектури

Перед початком безпосереднього впровадження нових функціональних елементів було здійснено систематичне ознайомлення з існуючою програмною архітектурою. Метою цього етапу було отримання вичерпних знань про системні залежності, логіку обробки просторових даних та існуючі архітектурні обмеження.

Для забезпечення ефективності та цілісності подальших модифікацій, процес ознайомлення включав кілька критично важливих кроків:

1. Ревізія технічної документації.

Детальне вивчення наявної документації, включно з файлами readme, для розуміння архітектурних рішень, використання бібліотек та існуючих інтерфейсів прикладного програмування (API).

2. Компіляція та верифікація функціональності.

Збірка існуючого програмного забезпечення для встановлення його базового функціонального стану та підтвердження працездатності поточних компонентів.

3. Експлоративне тестування.

Дослідження запущеного додатка з позиції кінцевого користувача, що дозволило ідентифікувати поточні патерни взаємодії та оцінити користувацький досвід (UX).

4. Аналіз архітектури коду.

Глибинний аналіз архітектурної схеми коду з метою визначення необхідних точок втручання та модифікацій для підтримки багатомодальної функціональності, яка не існує в поточному стані дизайну додатку (наприклад, інтеграція нових методів просторового введення).

На рисунку 3.1 відображено існуючі можливості системи, де підкреслюється, що планування та виконання ключових завдань наявне лише через автономний додаток для планшетних пристроїв. Це підтверджує критичну необхідність крос-платформної інтеграції інтерфейсу для забезпечення функціонування у просторових модальностях VR та AR.

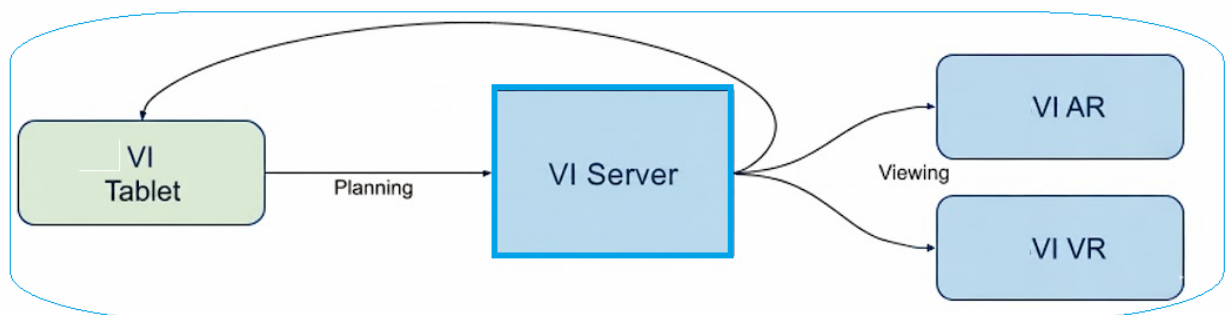


Рис. 3.1. Існуюча функціональність системи (VI), що демонструє, що планшет є єдиним методом для виконання завдань

Наступні підрозділи детально описують специфікації програмного та апаратного забезпечення, які були використані для проведення розробки, налагодження та фінальної реалізації.

3.1.2. Інтеграція прототипу та модифікація UI

Після аналізу архітектури, макети UI, створені в Adobe XD, були адаптовані та реалізовані у середовищі розробки цільової системи. Цей процес вимагав трансформації 2D-векторних активів у просторові компоненти з підтримкою 3D-взаємодії.

- Перетворення активів. Елементи UI були перенесені з векторного формату в текстури та моделі, оптимізовані для рендерингу у VR/AR.

- Реалізація взаємодії. Функціональність інтерактивності, змодельована в Adobe XD (клацання, наведення), була замінена на фізичні скрипти (наприклад, Unity C# або Unreal Blueprints) для підтримки просторових методів введення (ray casting, жести).

Для подальшої емпіричної валідації імплементованого UI було вбудовано механізми збору даних протоколу. Ці механізми були розроблені для реєстрації поведінкових метрик під час виконання користувачами визначених завдань.

Система фіксувала точний час початку та завершення кожного тестового завдання, а також послідовність взаємодій (наприклад, кліків, виборів меню, жестів) користувача.

Збиралися такі ключові показники, як час виконання завдання, кількість помилок (наприклад, невірних кліків або невірних шляхів) та коефіцієнт успішності. Ці дані були необхідні для проведення статистичного аналізу з метою порівняння ефективності використання нового, масштабованого UI у модальностях VR та AR.

3.2. Архітектурні вимоги та методи багатомодальної взаємодії

Для побудови функціонально ефективною та значущою системи шляхом інтеграції всіх необхідних компонентів, були сформульовані наступні ключові вимоги до носінної системи змішаної реальності (MR), як подано в таблиці 3.1.

Вимоги до системи змішаної реальності

Категорія вимоги	Ключова функціональність
Взаємодія	Віртуальний агент повинен розпізнавати об'єкти, на які сфокусовано погляд користувача, та ініціювати релевантну бесіду.
Виразність	Віртуальний агент має демонструвати емоції та жести тіла, які є когерентними зі змістом поточної розмови.
Масштабованість	Віртуальний агент повинен мати можливість позиціонування на необмеженій робочій області через використання надійних механізмів просторового прив'язування.

Загальна архітектурна схема, що об'єднує ці компоненти, представлена на рисунку 3.2.

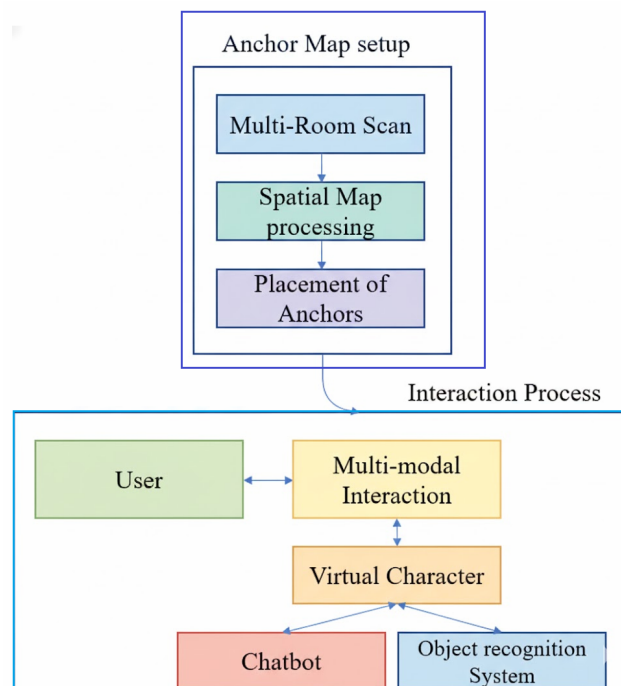


Рис. 3.2. Архітектура системи

Процес взаємодії був спроектований з метою забезпечення максимально природного та інтуїтивно зрозумілого методу спілкування з віртуальним персонажем. Ключова методологічна проблема полягала у створенні ефективного механізму для ініціювання першого запиту

користувача про невідомий об'єкт. Якщо користувач спочатку дивиться на агента, а потім швидко переводить погляд на об'єкт, може виникнути недостатній час для захоплення (фіксації) об'єкта системою розпізнавання. Ця проблема була вирішена шляхом інтеграції обмеженого набору голосових команд (наприклад, "Що це?", "Розкажи про це" тощо) для забезпечення негайної контекстної взаємодії.

На рисунку 3.3 детально зображено процес взаємодії як з наявними, так і з відсутніми просторовими прив'язками (якорями) в полі зору користувача.

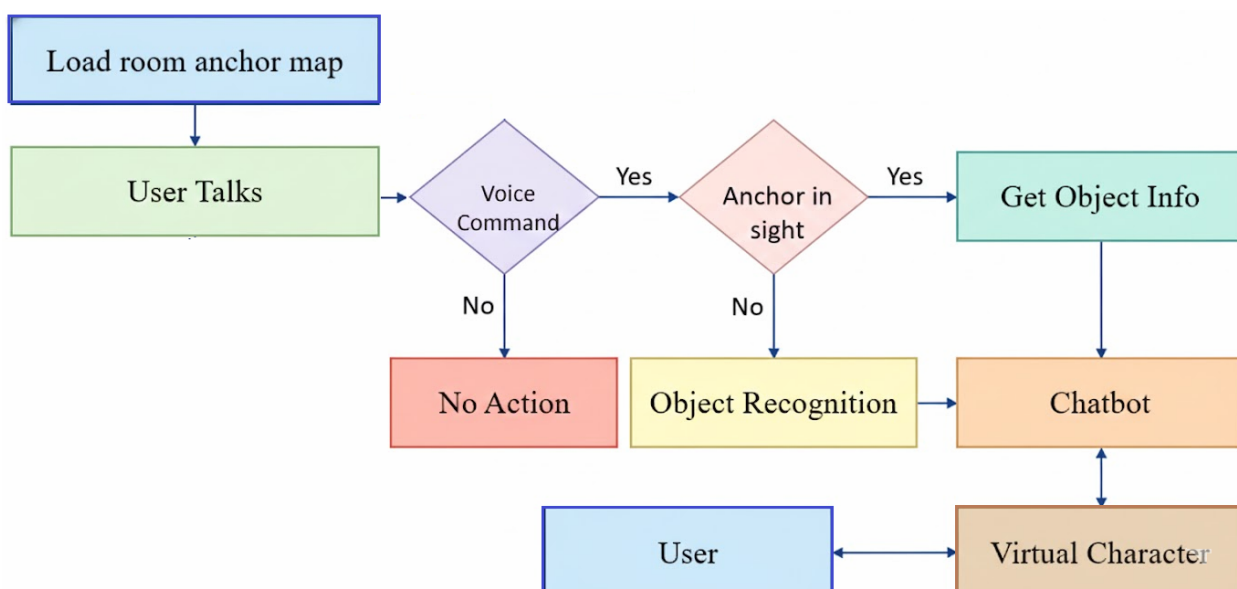


Рис. 3.3. Потік взаємодії користувача з віртуальним агентом у системі змішаної реальності

Рисунок 3.3 ілюструє потік взаємодії користувача з віртуальним агентом у системі змішаної реальності (MR), зосереджуючись на багатомодальному введенні (голос та погляд) та розпізнаванні об'єктів. Цей процес починається з ініціалізації просторового контексту.

1. Ініціалізація та введення користувача

Завантаження карти якорів кімнати (Load room anchor map) - процес починається із завантаження карти просторових якорів кімнати. Якорі — це постійні точки прив'язки у реальному світі, які дозволяють системі MR точно

позиціонувати віртуальні об'єкти (включаючи агента та інформаційні елементи) та розпізнавати оточення.

User Talks - система активується при виявленні мовлення користувача (голосового введення).

2. Аналіз голосової команди

Voice Command - система аналізує, чи є виявлене мовлення конкретною, попередньо визначеною голосовою командою (наприклад, "Що це?", "Розкажи про це"). Якщо мовлення не розпізнається як команда (наприклад, це була розмова з кимось іншим), система не виконує жодної дії.

3. Просторова валідація та розпізнавання

Anchor in sight. Якщо мовлення розпізнано як голосова команда (Yes), система перевіряє, чи знаходиться просторовий якір (пов'язаний з об'єктом) у поточному полі зору користувача.

Якщо ТАК, це означає, що користувач дивиться на вже ідентифікований або прив'язаний об'єкт. Система переходить до кроку отримання інформації про об'єкт (Get Object Info), пропускаючи загальне розпізнавання, оскільки об'єкт уже визначено якорем.

Якщо НІ, то це означає, що користувач дивиться на новий або не прив'язаний об'єкт. Система ініціює розпізнавання об'єктів.

4. Обробка інформації та взаємодія

Object Recognition. Система використовує свій менеджер зору для ідентифікації об'єкта в полі зору користувача та визначає його категорію або назву.

Chatbot. Інформація про ідентифікований об'єкт (отримана з Object Recognition) надсилається до інтелектуального чат-бота. Чат-бот генерує відповідь, використовуючи свою базу знань та додаючи параметри, як-от клас настрою.

Virtual Character. Відповідь чат-бота передається до віртуального персонажа. Персонаж візуалізує відповідь (синтез мовлення) та додає виразність (жести, міміку) відповідно до отриманих параметрів настрою.

Користувач отримує аудіо-візуальну відповідь від віртуального агента, завершуючи цикл взаємодії.

Таким чином, блок-схема описує механізм для багатомодального контекстуального запиту, що дозволяє агенту швидко реагувати на голосові команди, автоматично розпізнаючи об'єкт, на який сфокусовано погляд користувача, або використовуючи існуючі просторові прив'язки.

3.2.1. Багатомодальна взаємодія. Архітектура потоку діалогової взаємодії

Проведено значні дослідження щодо створення природних методів взаємодії людини з віртуальними агентами. Наприклад, у дослідженні [9] підкреслювалася роль мовлення та жестів рук у передачі прагматичної інформації. В роботі [23] запропонували підхід для інтерактивних віртуальних персонажів у MR-застосунках, де використання погляду та жестів підвищує рівень комфорту користувачів.

Згідно з [10], погляд та проксемічна поведінка можуть встановлювати та підтримувати рівень близькості, що стосується не лише взаємодії людина-людина, а й взаємодії людина-агент. Зокрема, було показано, що агенти з вищою візуальною близькістю викликали у учасників тенденцію відходити далі. Модель поведінки погляду для інтерактивного віртуального персонажа, розташованого в реальному світі, представлена у [22].

У даній системі застосовується комбінований підхід розпізнавання погляду та мовлення для взаємодії з віртуальним персонажем, використовуючи наявні API HoloLens. HoloLens Toolkit надає високорівневі функції для розпізнавання мовлення, такі як менеджер диктування, що підтримує обробку довгих речень.

Пропонований метод, проілюстрований на рисунку 3.4, не вимагає використання ключового слова активації. Рис. 3.4 ілюструє архітектуру керування діалоговим потоком між користувачем та віртуальним персонажем

у системі змішаної реальності (MR), сфокусовану на багатомодальній взаємодії за допомогою погляду та мовлення.

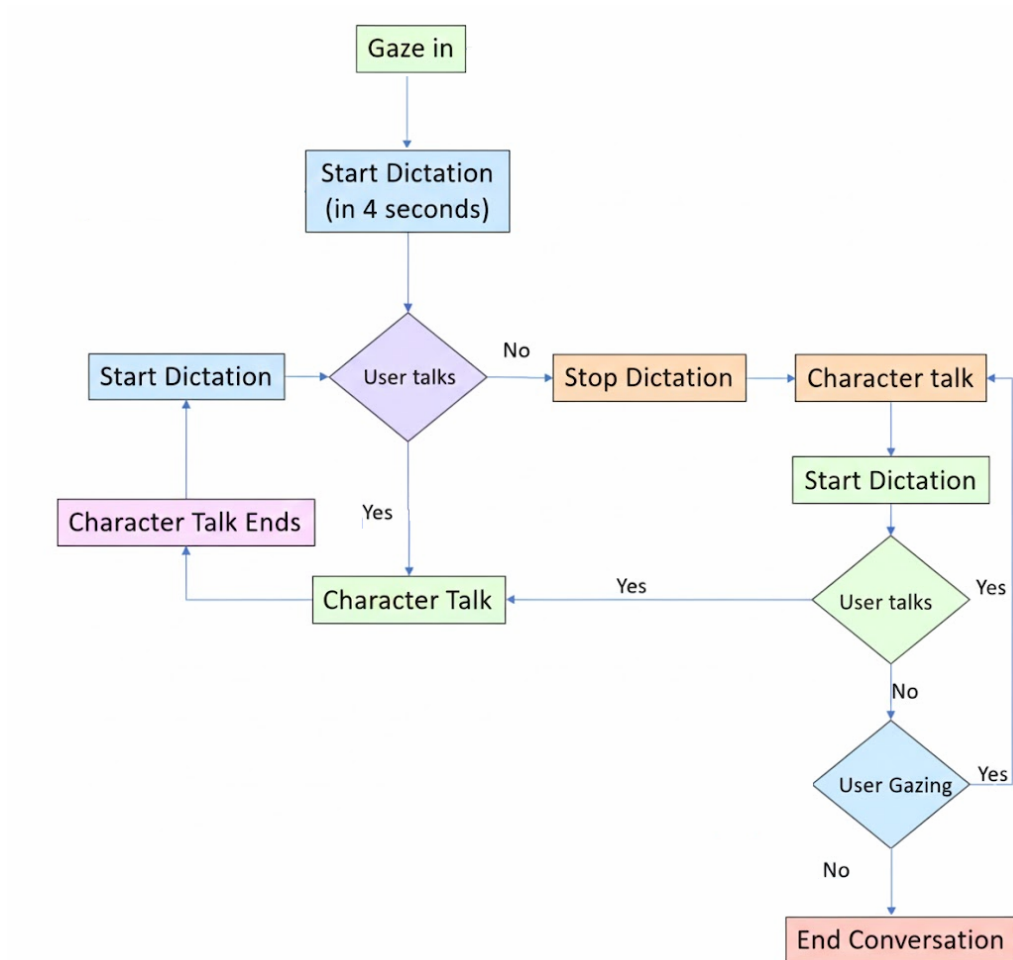


Рис. 3.4. Архітектура потоку діалогової взаємодії віртуального агента

Розглянемо цю архітектуру більш детально.

1. Ініціація розмови

- Gaze in. Діалог ініціюється, коли користувач фіксує свій погляд на віртуальному персонажі.

- Start Dictation (in 4 seconds). Після фіксації погляду система автоматично активує режим розпізнавання мовлення/диктування через заданий інтервал часу (4 секунди).

2. Обробка мовлення користувача

- User talks. Система очікує голосового введення від користувача.

Якщо ТАК, то розпізнане мовлення надсилається для обробки, і система переходить до Character Talk. Якщо користувач не говорить протягом часу очікування, система виконує такі дії:

- Stop Dictation - деактивується режим розпізнавання мовлення.
- Character talk - персонаж ініціює розмову (наприклад, "Привіт, чи можу я допомогти?").

3. Цикл відповіді персонажа та продовження діалогу

- Character talk. Персонаж надає відповідь або ініціює питання.

Після відповіді (сценарій 'Ні' на початку): Якщо персонаж ініціював розмову (сценарій 'Ні' у кроці 3), діалоговий потік намагається відновити режим диктування.

- Start Dictation. Система знову активує режим розпізнавання мовлення.
- User talks. Система очікує відповіді користувача.

Якщо ТАК, то діалог продовжується, і система повертається до відповіді персонажа (Character Talk). Якщо НІ, то перевіряється утримання уваги користувача.

- User Gazing. Якщо погляд користувача все ще сфокусований на персонажі (Yes), система повертається до відповіді персонажа (Character talk) для повторного запитання або очікування. Якщо НІ, то користувач відвів погляд, і розмова завершується (End Conversation).

Після відповіді (сценарій 'Так' на початку). Якщо персонаж відповів на запитання користувача (сценарій 'Так' у кроці 3), відбувається перехід до завершення розмови (End Conversation). Цей шлях завершення діалогу після відповіді є спрощеною архітектурою, де очікується, що новий цикл розмови повинен бути ініційований повторним поглядом (або голосовою командою).

4. End Conversation. Діалог припиняється, коли користувач відводить погляд, або після завершення одноразового циклу запитання-відповіді.

Ця архітектура демонструє використання багатомодального введення для створення природної, але керованої взаємодії:

- Використання фіксації погляду як неявного ключового слова для ініціації взаємодії.
- Система автоматично переходить у режим очікування відповіді (диктування) після ініціації, мінімізуючи необхідність ручної активації.
- Система чутлива до не лише мовлення, але й до уваги користувача (погляду) для визначення, чи слід продовжувати чи завершувати діалог.

3.2.2. Розпізнавання об'єктів поглядом

Система розпізнавання об'єктів виконує функцію "віртуальних очей" агента, забезпечуючи його обізнаність про навколишнє середовище. Може бути використана будь-яка сучасна або спеціально навчена система розпізнавання об'єктів чи класифікації зображень, за умови, що вона підтримує обробку веб-запитів.

Модуль менеджера зору відповідає за завдання розпізнавання об'єктів і потребує лише дійсного API-кінцевого пункту. Користувач має можливість створити власний API-кінцевий пункт або використовувати існуючий. Менеджер зору розроблений з урахуванням гнучкості, що дозволяє підключення до системи на пристрої (on-device processing). Завдяки модульному проектуванню, система здатна обробляти будь-яку кількість об'єктів, якщо чат-бот має відповідну базу знань.

3.2.3. Архітектура інтелектуального віртуального агента

Для забезпечення ефективної взаємодії віртуального персонажа з користувачем необхідна база знань та відповідна програма для надання необхідної інформації, відома як чат-бот. Чат-боти відіграють важливу роль у взаємодії людини з комп'ютером і використовуються для автоматизації обслуговування.

У пропонованій системі чат-бот виконує дві ключові функції:

- Надання предметно-орієнтованої інформації.
- Обробка загальних варіантів бесіди.

Чат-бот оснащений системою аналізу настроїв (sentiment analysis), яка аналізує відповідь і додає клас настрою та рівень настрою, як показано на рисунку 3.5. Цей чат-бот розгорнутий у хмарному середовищі та доступний через веб-запити.

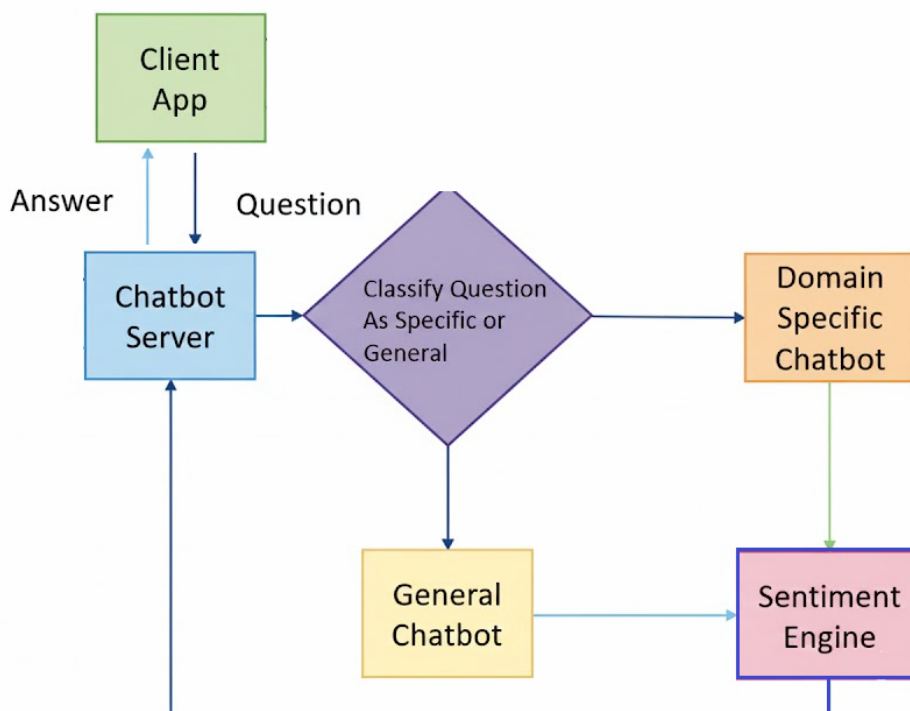


Рис. 3.5. Архітектура чат-боту

Вхідні параметри чат-бота мають структуру (запит, об'єкт):

- Запит може бути питанням, коментарем або відповіддю на питання агента.

Об'єкт — це будь-який виявлений віртуальний або реальний об'єкт, про який формується запит.

Запит може бути загальним (наприклад, про самого персонажа) або специфічним (про об'єкти в оточенні). Передумовою для специфічного запиту про об'єкт є необхідність першого погляду користувача на об'єкт, що запускає компонент розпізнавання об'єктів та передає інформацію до чат-бота. Наступні запити про той самий об'єкт не вимагають повторного наведення погляду.

Відповідь на запит має структуру: Відповідь, Клас настрою, Рівень настрою. Класи та рівні настрою використовуються для коректної анімації персонажа. Користувач може ставити як загальні, так і специфічні запити в рамках однієї бесіди та в довільному порядку. Такий стиль бесіди імітує взаємодію людина-людина, що значно підвищує залученість та якість користувацького досвіду.

3.3. Візуальна експресія та анімаційний контроль віртуального персонажа

Для досягнення високого рівня правдоподібності та реалізму віртуальний персонаж реалізований як деталізована 3D-модель, що вимагає імплементації багатих анімацій тіла та емоцій обличчя. Ці анімації поділяються на вербальні та невербальні дії. З огляду на необхідність персонажа рухатися та реагувати на вербальні стимули, система вимагає великого репертуару доступних анімацій.

Наша система містить 30 різних анімаційних послідовностей для забезпечення різноманітних вербальних та невербальних дій.

Система підтримує чотири основні класи емоцій: радість, гнів, сум та страх. Кожна емоція має три рівні інтенсивності: високий, середній та низький, як це схематично показано на рисунку 3.6.

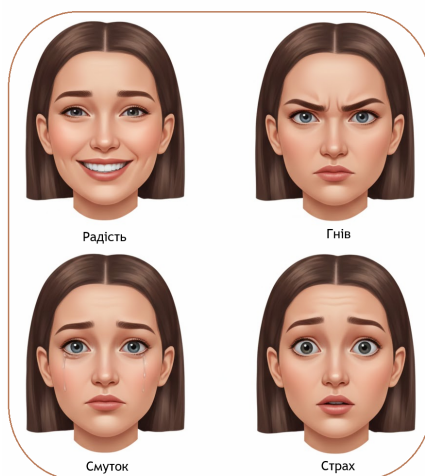


Рис. 3.6. Представлення емоцій віртуального персонажа

Анімації тіла та емоції обличчя застосовуються відповідно до семантичного змісту мовлення, яке генерується з тексту за допомогою системи перетворення тексту на мовлення (Text-to-Speech, TTS).

3.3.1. Архітектура синхронізації анімацій

Архітектура для побудови послідовності анімацій представлена на рисунку 3.7. Ця архітектура вимагає наявності таблиці відображення (mapping table), яка асоціює певний текст із відповідними анімаційними послідовностями.

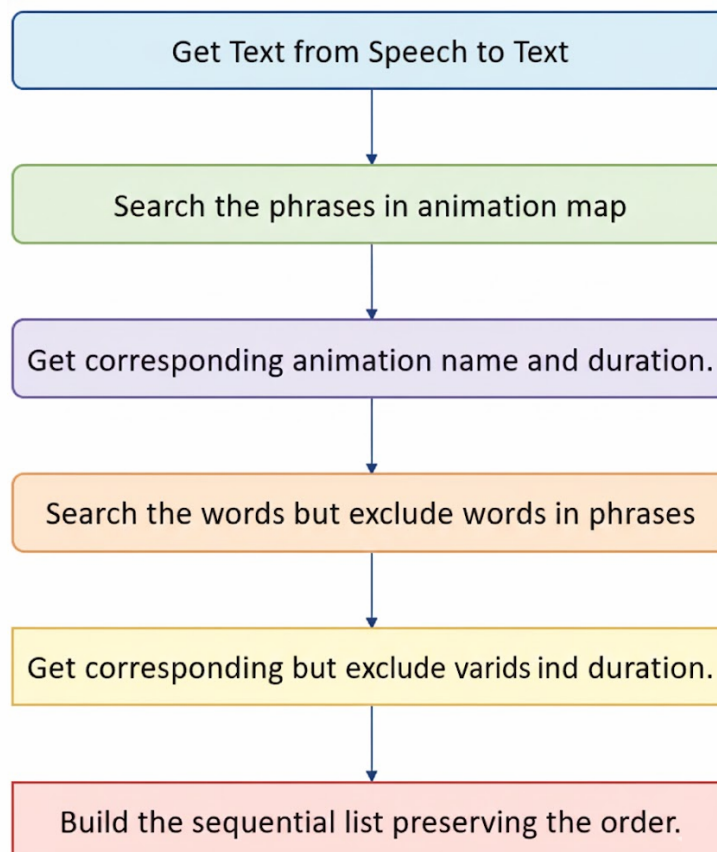


Рис. 3.7. Архітектура побудови послідовності анімацій віртуального агента

Рис. 3.7 ілюструє архітектуру керування анімаціями тіла віртуального персонажа, яка призначена для синхронізації невербальної комунікації (жестів) з вербальним мовленням. Ключова мета архітектури — перетворити

текстовий вивід системи на послідовний та осмислений набір анімаційних команд, зберігаючи при цьому порядок слів та фраз.

1. Введення тексту

Процес починається з отримання текстового рядка (Text String), який був згенерований системою Speech-to-Text (STT) на основі голосового введення користувача або тексту, згенерованого чат-ботом.

2. Пошук фраз у карті анімацій (Search the phrases in animation map). Система спочатку виконує пошук попередньо визначених ключових фраз у спеціалізованій таблиці відображення анімацій (animation map), створеній експертами. Фрази мають пріоритет перед окремими словами, оскільки вони несуть більш повне семантичне значення (наприклад, "дякую вам" або "я не знаю").

Отримати відповідну назву та тривалість анімації (Get corresponding animation name and duration) - для знайдених фраз із карти витягуються назва анімації (яка відповідає жесту) та її часова тривалість.

3. Пошук окремих слів (низький рівень)

Пошук Слів, але виключити слова у фразах (Search the words but exclude words in phrases). На наступному етапі система виконує пошук анімацій для окремих слів, які залишилися в текстовому рядку і не були частиною знайдених раніше фраз. Це запобігає дублюванню та гарантує, що анімації призначаються лише до унікальних сегментів тексту.

Отримати відповідну назву та тривалість анімації (Get corresponding animation name and duration). Для цих окремих слів також витягуються назви та тривалість відповідних анімацій (наприклад, жести, що відповідають словам "великий", "багато" тощо).

4. Побудова фінальної послідовності

Побудувати послідовний список, зберігаючи порядок (Build the sequential list preserving the order). Фінальним кроком є компіляція всіх витягнутих анімаційних команд (для фраз і для слів) в єдиний послідовний список. При цьому критично важливо зберегти оригінальний порядок слів і

фраз у реченні, щоб забезпечити синхронність жесту та відповідного слова в мовленні.

Цей послідовний список потім використовується системою керування анімаціями для паралельного відтворення жестів із синтезованим мовленням.

Карта анімацій створюється експертами і містить відображення анімацій як для повних фраз, так і для окремих слів. Ці дані використовуються конструктором анімацій (animation builder) для формування фінальної послідовності. Емоції обличчя визначаються на основі тексту за допомогою системи аналізу настроїв (sentiment analysis), інтегрованої в чат-бот, і застосовуються протягом усього часу тривалості мовлення. Синхронізація мовлення з рухами губ персонажа (ліпсинк) здійснюється шляхом перетворення тексту на послідовність фонем. Для цього використовується спеціалізована карта фонем (phoneme map).

3.3.2. Архітектура обробки запиту та генерації відповіді віртуального персонажа

Застосування мовлення, анімації тіла та емоцій обличчя здійснюється паралельно, що гарантує безшовну та реалістичну візуально-аудіальну репрезентацію агента (рисунок 3.8).

Рисунок 3.8 ілюструє архітектуру обробки запиту та генерації багатомодальної відповіді віртуального персонажа. Ця архітектура показує, як єдиний запит користувача трансформується у синхронізований вивід (мовлення, жести та міміка) віртуального агента.

Розглянемо дану архітектуру дещо детальніше.

1. Обробка запиту чат-ботом

- Запит (Query). Процес ініціюється, коли сформований запит користувача (отриманий після обробки мовлення та розпізнавання об'єкта) надходить до системи.

- Чат-бот (Chatbot). Запит обробляється інтелектуальним чат-ботом, який використовує свою предметно-орієнтовану базу знань.

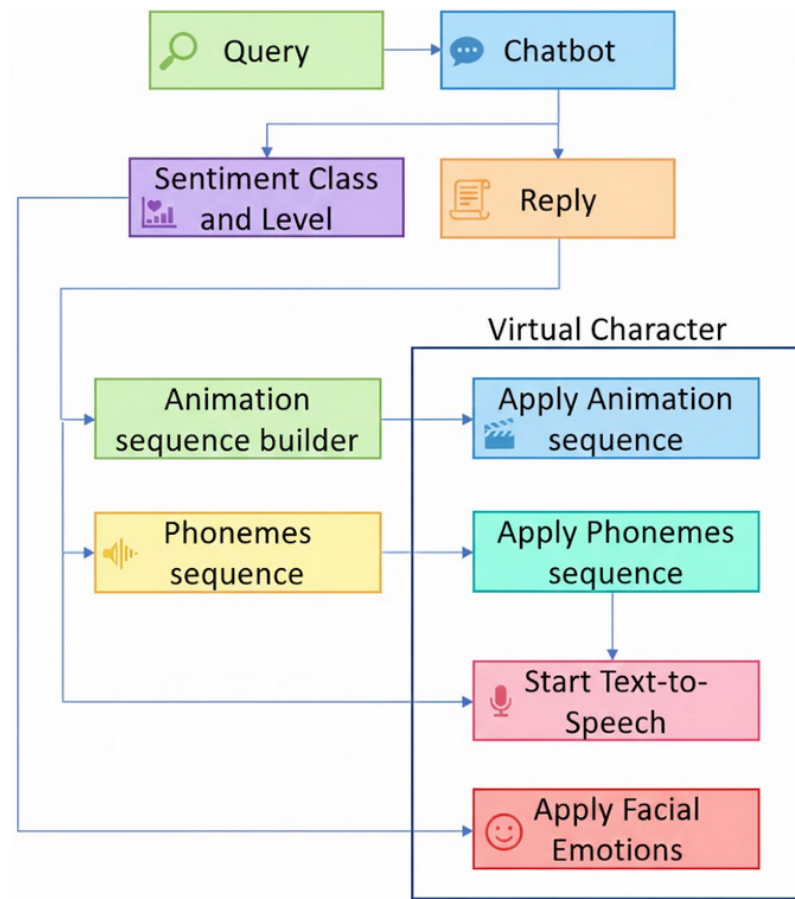


Рис. 3.8. Архітектура обробки запиту та генерації відповіді віртуального персонажа

Чат-бот генерує два паралельні вихідні параметри:

- Клас та рівень настрою (Sentiment Class and Level). Система аналізу настроїв, інтегрована в чат-бот, визначає емоційний тон відповіді (наприклад, радість, середній рівень).

- Відповідь (Reply). Власне текстова відповідь, яку має озвучити персонаж.

2. Генерація мультимодального виводу

Результати роботи чат-бота запускають три паралельні процеси, які готують вивід для віртуального персонажа:

- Animation sequence builder - текстова відповідь (Reply) передається до конструктора, який використовує карту відображення слів/фраз на анімації (як описано в попередніх архітектурах). Результатом є послідовність анімацій тіла, синхронізована з мовленням.

Phonemes sequence - текстова відповідь (Reply) також трансформується у послідовність фонем. Ця послідовність необхідна для точної синхронізації рухів губ (ліпсінк).

3. Синхронізоване відтворення віртуальним персонажем

Всі підготовлені компоненти інтегруються та відтворюються одночасно модулем Virtual Character:

- Apply Animation sequence. Відбувається відтворення жестів тіла відповідно до згенерованої послідовності.

- Apply Phonemes sequence. Активуються рухи губ (віземи), синхронізовані з вимовою.

- Start Text-to-Speech. Запускається генерація звукового мовлення на основі текстової відповіді (Reply).

- Apply Facial Emotions. На обличчя персонажа накладається міміка, що відповідає отриманому класу та рівню настрою.

Таким чином, ця архітектура забезпечує синхронне та когерентне багатомодальне представлення відповіді, підвищуючи реалізм та ефективність комунікації віртуального агента.

3.4. Забезпечення масштабованості та роботи віртуального агента

Для забезпечення масштабованості та роботи віртуального агента на необмеженій робочій області використовується вдосконалений механізм керування просторовими прив'язками.

HoloLens надає вбудований механізм, відомий як просторові якорі (Spatial Anchors), який дозволяє відстежувати статичні об'єкти в реальному світі.

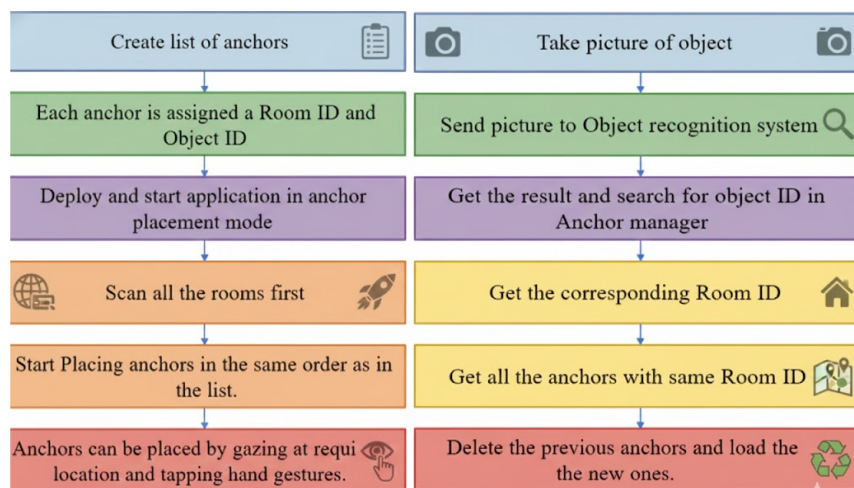
Шляхом інтеграції розпізнавання зображень та просторових якорів було створено процес, який дозволяє відображати необмежену робочу область та зберігати інформацію про реальні об'єкти, до яких може звертатися віртуальний агент. Цей підхід значно мінімізує необхідність

виконання ресурсоємної операції розпізнавання об'єктів для кожного об'єкта під час сесії.

Процес початкового розміщення якорів (Anchor Placement) деталізовано на рисунку 3.9 а).

Оскільки нативна функціональність HoloLens не містить механізму для надійного завантаження якорів, специфічних для конкретної кімнати (існує ризик завантаження нерелевантних якорів для ідентичних кімнат), була розроблена додаткова процедура.

Для уникнення проблеми некоректного завантаження, процес управління якорями використовує розпізнавання об'єктів для ідентифікації та завантаження якорів, специфічних для поточного просторового середовища. Цей вдосконалений процес показано на рисунку 3.9 б).



а) Процес початкового розміщення якорів б) Вдосконалений процес

Рис. 3.9. Процес розміщення якорів

3.5. Архітектура та технологічна платформа системи змішаної реальності

Розроблена платформа базується на інтеграції таких ключових технологічних компонентів: Unity3D, Microsoft HoloLens, HoloLens Toolkit та хмарна обчислювальна платформа.

Microsoft HoloLens є просторовим голографічним пристроєм змішаної реальності (MR). Цей пристрій використовується для сканування фізичного оточення та створення 3D-моделі простору, яка слугує основою для відстеження положення користувача, віртуальних об'єктів та інтерактивного віртуального персонажа. HoloLens підтримує Unity3D як основну платформу розробки.

HoloLens Toolkit (Mixed Reality Toolkit, MRTK) використовується для спрощення початкового налаштування проекту та параметрів. Він також надає численні допоміжні функції високого рівня, що стосуються просторового картування, відстеження погляду та розпізнавання мовлення.

HoloLens є автономним пристроєм з обмеженими обчислювальними ресурсами. Запуск ресурсомістких процесів, таких як чат-бот та система розпізнавання об'єктів, безпосередньо на пристрої призводить до зниження продуктивності та компромісів у якості користувацького досвіду (UX).

Для вирішення цього архітектурного обмеження, наша система реалізує розподілену обчислювальну модель. Доступ до чат-бота та системи розпізнавання об'єктів здійснюється через хмарну обчислювальну платформу. Зв'язок між пристроєм та хмарними сервісами забезпечується за допомогою менеджера модулів, що дозволяє ефективно обробляти складні завдання поза пристроєм.

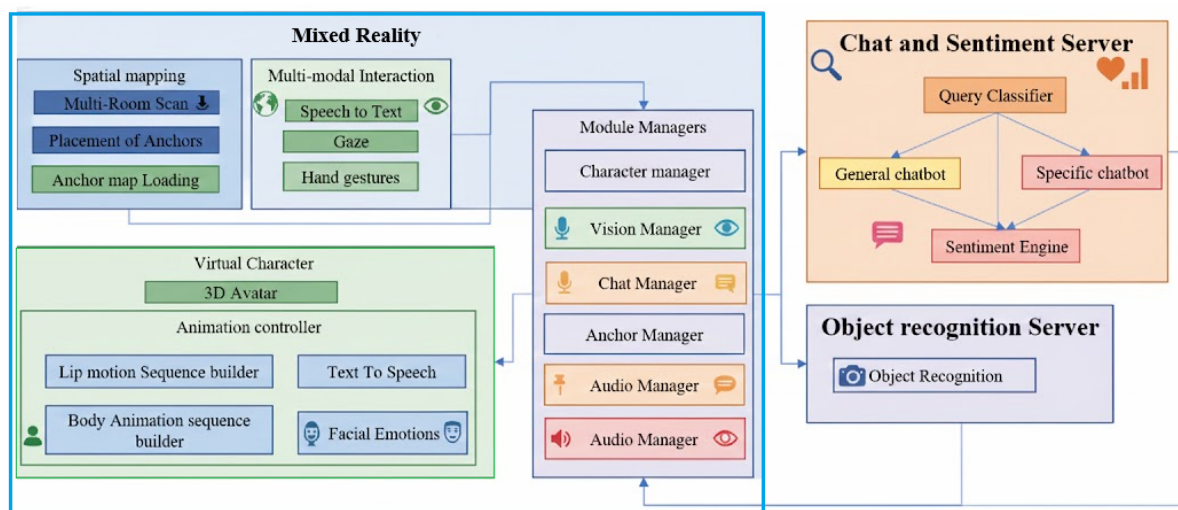


Рис. 3.10. Архітектура мультимодальної системи віртуального агента

Рисунок 3.10 ілюструє загальну архітектуру фреймворку для багатомодальної системи змішаної реальності з віртуальним агентом, яка складається з чотирьох основних, взаємопов'язаних модулів.

1. Модуль Mixed Reality (клієнтський інтерфейс та введення)

Цей модуль представляє функціональність на пристрої, відповідальну за введення та просторову обізнаність:

- Просторове картування (spatial mapping) включає сканування кількох кімнат (Multi-Room Scan), розміщення просторових якорів (Placement of Anchors) та завантаження карти якорів (Anchor map Loading) для орієнтації віртуальних об'єктів.

- Мультимодальна взаємодія (multi-modal interaction) відповідає за збір даних введення від користувача, включаючи перетворення мовлення на текст (Speech to Text), відстеження погляду (Gaze) та жести рук (Hand gestures).

2. Модулі керування

Цей набір модулів є ядром системи, що керує потоками даних між клієнтом (HoloLens) та хмарними серверами:

- Включає керівників для персонажа (Character manager), зору (Vision Manager), чату (Chat Manager), якорів (Anchor Manager), аудіо (Audio Manager) та погляду (Gaze Manager). Вони організують та маршрутизують запити до відповідних компонентів.

3. Сервер обробки чату та настроїв (chat and sentiment server)

Цей хмарний модуль обробляє лінгвістичні запити:

- Класифікатор запитів (query classifier) сортує вхідні запити на загальні (General chatbot) та специфічні (Specific chatbot).

- Аналіз настроїв (sentiment engine) обробляє вихідну відповідь від чат-ботів для визначення класу та рівня настрою, необхідних для емоційної виразності персонажа.

4. Сервер розпізнавання об'єктів (Object recognition Server)

Окремий хмарний модуль для ресурсомістких завдань ідентифікації, що містить систему розпізнавання об'єктів (Object Recognition), яка

використовується менеджером зору для ідентифікації об'єктів у полі зору користувача, коли інформація про них не може бути отримана через якорі.

5. Модуль віртуального персонажа (virtual character)

Цей модуль відповідає за фінальний, багатомодальний вивід для користувача:

- включає 3D аватар та контролер анімацій (Animation controller).
- використовує Text to Speech для генерації мовлення.
- застосовує емоції обличчя (Facial Emotions) на основі даних про настрій, отриманих із хмари.

Отже, архітектура демонструє розподілену систему, де HoloLens збирає просторові та багатомодальні дані, які обробляються менеджерами, а обчислювально складні функції (чат та розпізнавання) делегуються хмарним серверам. Фінальний синтез відбувається на модулі віртуального персонажа.

Висновки до розділу

Третій розділ присвячено безпосередній реалізації методів і моделей масштабування мультимодальних інтерфейсів, що дозволило трансформувати концептуальні напрацювання у функціональну XR-систему. На основі аналізу й декомпозиції існуючої архітектури було визначено основні модулі системи та їхню роль у забезпеченні стабільної й масштабованої взаємодії. Інтеграція прототипу, розробленого у попередньому розділі, дозволила перевірити відповідність концептуального дизайну вимогам реальної системи та виконати необхідну модифікацію UI. Окрему увагу приділено архітектурі багатомодальної взаємодії, де було реалізовано поєднання голосових команд, жестових сигналів, просторової навігації та контролю за допомогою погляду. У роботі представлено детальну схему діалогового потоку, яка забезпечує природність комунікації між користувачем і системою, а також описано механізми розпізнавання об'єктів поглядом, що підвищують точність взаємодії. Розробка архітектури

інтелектуального агента та його анімаційної експресії дала змогу поєднати когнітивну модель поведінки з візуальною і голосовою реакцією персонажа, створюючи цілісне мультимодальне середовище. У розділі також було описано механізми синхронізації анімацій, що забезпечують узгодженість між мовленням агента, його позою та мімікою, формуючи переконливу і реалістичну взаємодію.

ВИСНОВКИ

Магістерська робота, присвячена дослідженню методів і моделей масштабування мультимодальних інтерфейсів у віртуальній та доповненій реальності, продемонструвала комплексний підхід до аналізу, проектування та практичної реалізації XR-систем, орієнтованих на взаємодію користувача з інтелектуальними агентами та складними тривимірними середовищами. Отримані результати охоплюють декілька взаємопов'язаних аспектів: теоретичні засади масштабованості XR-інтерфейсів, методологічні засоби швидкого прототипування, архітектурні рішення мультимодальної взаємодії та практичну реалізацію моделей, що забезпечують стабільність, адаптивність і розширюваність систем змішаної реальності.

У першому розділі проведено системне дослідження предметної області масштабованих XR-інтерфейсів. Розкрито особливості крос-платформенності та визначено ключові параметри масштабованості, які впливають на ефективність користувацької взаємодії. Континуум віртуальності, а також класифікація технологій розширеної реальності дали змогу чітко розмежувати вимоги до інтерфейсів різних типів XR-середовищ. Особлива увага була приділена ролі швидкого прототипування та систематичному аналізу доступних інструментів для створення високоточної взаємодії та попереднього моделювання користувацьких сценаріїв. Дослідження продемонструвало наявність прогалів у методиках масштабування UI/UX у хмарних, розподілених та багатомодальних XR-системах, що зумовило потребу в розробці нових принципів дизайну, орієнтованих на адаптивність, природність взаємодії, мінімізацію когнітивного навантаження та уніфікацію мультимодальних каналів.

Другий розділ був спрямований на розробку методології проектування XR-інтерфейсів із використанням швидкого прототипування. На основі побудови персонажів користувачів, афініті-діаграмування та створення початкових макетів сформовано послідовний процес еволюції інтерфейсу від

концептуальних ескізів до інтерактивних прототипів. У роботі обґрунтовано доцільність використання Adobe XD та інших сучасних інструментів для моделювання XR-прототипів, що забезпечують візуальну відповідність та мінімізацію витрат на ранніх етапах розроблення. Окремий внесок становить представлена методологія побудови інтелектуальних віртуальних агентів, яка інтегрує принципи когнітивного дизайну, діалогової взаємодії та поведінкових моделей у середовищах змішаної реальності. Це дозволило сформувати основу для подальшої реалізації мультимодальних агентів, здатних до природної комунікації через голос, жест, погляд та інтерактивні елементи.

Третій розділ містить практичну імплементацію методів і моделей масштабування мультимодальних XR-інтерфейсів. Проведений аналіз архітектури дозволив виконати її декомпозицію та створити модульну структуру, орієнтовану на інтеграцію UI, агентного шару та системи рендерингу. Було реалізовано архітектурні рішення для багатомодальної взаємодії, включно з моделлю діалогового потоку, розпізнаванням об'єктів поглядом та впровадженням інтелектуального віртуального агента. Значним технічним результатом роботи є розробка архітектури синхронізації анімацій, яка забезпечує узгодженість невербальної експресії та голосових реплік персонажа. Крім того, описано механізми підтримання масштабованості агентної системи, що дозволяють розширювати її функціональність без повної зміни архітектурного ядра та забезпечують стабільність при різних рівнях навантаження. Завершальний етап включає створення технологічної платформи змішаної реальності, яка інтегрує всі представлені методи та забезпечує єдине середовище для розгортання, тестування та масштабування розробленої системи.

Узагальнюючи результати дослідження, можна стверджувати, що магістерська робота виконала поставлену мету та вирішила комплекс завдань, спрямованих на підвищення ефективності проектування та реалізації масштабованих мультимодальних інтерфейсів для XR-середовищ.

Запропоновані моделі та методи демонструють важливість мультимодальності, адаптивності та інтелектуальної поведінки агентів для створення сучасних XR-систем. Розроблена архітектура, методологія прототипування та практичні програмні рішення становлять значний внесок у розвиток проектування інтерфейсів змішаної реальності та відкривають перспективи для подальших досліджень у напрямку інтеграції штучного інтелекту, автономних агентів та масштабованих XR-платформ.

ПЕРЕЛІК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Open Access proceedings Journal of Physics: Conference series - Aladin_2020_IOP_Conf._Ser.:_Mater._Sci._Eng._979_012004.pdf. - <https://iopscience.iop.org/article/10.1088/1757-899X/979/1/012004/pdf>
2. Siti Soleha - 6824 - A_Review_of_Multimodal_Interaction_Technique.pdf. - https://www.researchgate.net/profile/Lam-Meng-Chun/publication/328919_A_Review_of_Multimodal_Interaction_Technique_in_Augmented_Reality_Environment/links/5bf2d26692851c6b27cacb74/A-Review-of-Multimodal-Interaction-Technique-in-Augmented-Reality-Environment.pdf
3. Integrating Large Language Models with Multimodal Virtual Reality Interfaces to Support Collaborative Human-Robot Construction Work - <https://arxiv.org/pdf/2404.03498>
4. Milgram, P., & Kishino, F. “A taxonomy of mixed reality visual displays.” IEICE Transactions on Information Systems, Tokyo, Japan: IPCJ/IEICE, 1994, pp. 1321–1329.
5. Azuma, R. T. “A survey of augmented reality.” Presence: Teleoperators and Virtual Environments, Cambridge, MA: MIT Press, 1997, Vol. 6, No. 4, pp. 355–385.
6. Bowman, D. A., Kruijff, E., LaViola, J. J., & Poupyrev, I. “3D User Interfaces: Theory and Practice.” Boston, MA: Addison-Wesley, 2004, pp. 1–312.
7. Azuma, R., & Furmanski, C. “Evaluating display performance for augmented reality.” IEEE Computer Graphics and Applications, Los Alamitos, CA: IEEE Computer Society, 2003, Vol. 23, No. 6, pp. 54–63.
8. Billingham, M., Kato, H., & Poupyrev, I. “The MagicBook: moving seamlessly between reality and virtuality.” IEEE Computer Graphics and Applications, Los Alamitos, CA: IEEE Computer Society, 2001, Vol. 21, No. 3, pp. 6–8.

9. Azuma, R., Baillot, Y., Behringer, R., Feiner, S., & Julier, S. "Recent advances in augmented reality." *IEEE Computer Graphics and Applications*, Los Alamitos, CA: IEEE Computer Society, 2001, Vol. 21, No. 6, pp. 34–47.
10. Oviatt, S. "Multimodal interfaces." *Communications of the ACM*, New York, NY: ACM Press, 1999, Vol. 42, No. 2, pp. 31–37.
11. Oviatt, S. "Ten myths of multimodal interaction." *Communications of the ACM*, New York, NY: ACM Press, 2006, Vol. 50, No. 3, pp. 24–26.
12. Jacob, R. J. K., Girouard, A., Hirshfield, L. M., Horn, M. S., Shaer, O., Solovey, E. T., & Zigelbaum, J. "Reality-based interaction: a framework for post-WIMP interfaces." *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI)*, New York, NY: ACM Press, 2008, pp. 201–210.
13. Leiva, G., Nguyen, C., Habib Kazi, R., & Asente, P. "PRONTO: Rapid augmented reality video prototyping using sketches and enaction." *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, Honolulu, HI, USA: ACM Press, 2020, pp. 1–14.
14. Leiva, G., et al. "Pronto: Rapid augmented reality video prototyping." *CHI 2020 Proceedings*, New York, NY: ACM Press, 2020, pp. 1–13.
15. Freitas, G., & Gerber, E. "A systematic review of rapid prototyping tools for augmented reality." *Proceedings of the International Symposium on Mixed and Augmented Reality (ISMAR) Workshop*, New York, NY: IEEE/ACM, 2019, pp. 45–60.
16. Rapido Project Team. "Rapido: Prototyping interactive AR experiences through programming by demonstration." *Proceedings of UIST / CHI Workshop on Rapid Prototyping in XR*, New York, NY: ACM, 2021, pp. 7–18.
17. Leiva, G., Nguyen, C., Habib Kazi, R., & Asente, P. "Pronto: Rapid augmented reality video prototyping using sketches and enaction." *CHI 2020*, Honolulu, HI, New York, NY: ACM Press, 2020, pp. 1–12.

- 18.Chen, X., Li, Y., & Zhang, H. “Gaze-based interaction intention recognition in virtual reality.” *Electronics*, Basel, Switzerland: MDPI, 2022, Vol. 11, No. 10, pp. 1–18.
- 19.Bektaş, K., & Sormaz, D. “Gaze-enabled activity recognition for augmented reality.” *Pattern Recognition Letters*, Amsterdam, Netherlands: Elsevier, 2024, Vol. 178, pp. 20–31.
- 20.Yuan, Z., & Höllerer, T. “MEinVR: Multimodal interaction techniques in immersive molecular visualization.” *Journal on Computing and Cultural Heritage*, Amsterdam, Netherlands: Elsevier, 2023, Vol. 16, pp. 1–15.
- 21.Dritsas, E., & Stylianou, A. “Multimodal interaction, interfaces, and communication: survey and future directions.” *Multimodal Technologies and Interaction*, Basel, Switzerland: MDPI, 2025, Vol. 9, No. 1, pp. 1–29.
- 22.Börsting, I., & Broll, W. “Design patterns for mobile augmented reality user interfaces.” *Information*, Basel, Switzerland: MDPI, 2022, Vol. 13, No. 4, pp. 1–22.
- 23.Chu, C. H., & Chen, P. “Augmented reality user interface design and experimental evaluation for human–robot collaborative assembly.” *Applied Ergonomics*, Amsterdam, Netherlands: Elsevier, 2023, Vol. 104, pp. 1–12.
- 24.Smith, J., & Johnson, L. “Evaluating the scalability of non-preferred hand mode switching in augmented reality.” *Proceedings of the ACM Symposium on User Interface Software and Technology (UIST)*, New York, NY: ACM Press, 2020, pp. 123–134.
- 25.Kato, H., Billingham, M., & Poupyrev, I. “Tangible interaction in augmented reality systems.” *Proceedings of International Conference on Tangible and Embedded Interaction (TEI)*, New York, NY: ACM Press, 2005, pp. 89–96.
- 26.Azuma, R., et al. “Issues in evaluating augmented reality systems.” *ISMAR Proceedings*, Los Alamitos, CA: IEEE Computer Society, 1999, pp. 1–10.

27. Bowman, D. A., et al. "Design guidelines for virtual environments: interaction and presence." *Virtual Reality*, London, UK: Springer, 2001, Vol. 5, No. 2, pp. 2–17.
28. Steed, A., & Slater, M. "A framework for the evaluation of immersive virtual environments." *Proceedings of IEEE Virtual Reality Conference*, Los Alamitos, CA: IEEE, 2012, pp. 49–56.
29. Kording, K., & Wolpert, D. "The relevance of motor control to interaction design in VR." *Journal of Neural Engineering*, Bristol, UK: IOP Publishing, 2010, Vol. 7, No. 4, pp. 1–12.
30. Dey, A. K. "Understanding and using context." *Personal and Ubiquitous Computing*, London, UK: Springer, 2001, Vol. 5, No. 1, pp. 4–7.
31. Calvo, R. A., & Peters, D. "Designing adaptive multimodal systems for cognitive load reduction." *International Journal of Human-Computer Studies*, Amsterdam, Netherlands: Elsevier, 2018, Vol. 112, pp. 1–15.
32. Khamis, M., & Grossman, T. "Designing multimodal input for immersive analytics." *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT)*, New York, NY: ACM Press, 2021, Vol. 5, No. 2, pp. 1–20.
33. Roussel, N., & Venkatesh, A. "Synchronization of speech and facial animation for believable virtual agents." *Computer Animation and Virtual Worlds*, Chichester, UK: John Wiley & Sons, 2016, Vol. 27, No. 3–4, pp. 241–253.
34. Cassell, J., Sullivan, J., Prevost, S., & Churchill, E. "Embodied Conversational Agents." Cambridge, MA: MIT Press, 2000, pp. 1–312.
35. Traum, D., & Larsson, S. "Dialogue modeling for multimodal conversational agents." *Proceedings of the International Conference on Multimodal Interaction (ICMI)*, New York, NY: ACM Press, 2019, pp. 55–64.

36. Kopp, S., Krämer, N., & Wachsmuth, I. "Towards a common framework of multimodal behavior generation." In: *Multimodal Behavior: A Model and its Implementation*, Berlin, Germany: Springer, 2008, pp. 1–19.
37. Cassell, J. "Towards a model of technology and literacy in multimodal interfaces." *Journal of Computer-Mediated Communication*, Oxford, UK: Oxford University Press, 2004, Vol. 9, No. 1, pp. 1–23.
38. Hwang, S., & Kaufmann, H. "Modular architectures for scalable mixed reality systems." *IEEE Transactions on Visualization and Computer Graphics*, Los Alamitos, CA: IEEE, 2017, Vol. 23, No. 11, pp. 2385–2396.
39. Piumsomboon, T., Clark, A., & Billingham, M. "User-centric adaptation of AR interfaces for different form factors." *Proceedings of ISMAR*, Los Alamitos, CA: IEEE, 2019, pp. 100–109.
40. Kim, J., & Lee, S. "Cloud-assisted rendering and streaming for scalable VR experiences." *IEEE Communications Magazine*, New York, NY: IEEE, 2018, Vol. 56, No. 2, pp. 34–41.
41. Boring, S., & Baur, D. "No-code and low-code approaches for XR prototyping: comparative evaluation." *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*, New York, NY: ACM Press, 2021, pp. 1–12.