Developing Gaze-Based Map Interactions in Mixed Reality Devices

Master's thesis presentation

Kurumbayeva Nargiz

Chair Professor: Prof. Dr.-Ing. Liqiu Meng

Supervisor: Dr.-Ing. Christian Murphy

Reviewer: MSc. Wangshu Wang

Introduction

Motivation

- Eye-tracking is becoming pervasive in various research fields (Mardanbegi et al., 2016) and entering the mass market (HoloLens 2 from Microsoft).
- … gaze-supported interaction can contribute to cartographic applications.
- … can facilitate the hand gestures combined with the gaze control as suggested by Schweigert et al. (2019).
- … can offer unusual possibilities for differently-abled people and propose new opportunities in human-computer interaction (Piotrowski & Nowosielski, 2020).
- … can possibly allow for an easy, natural, and fast way of interacting in mixed reality devices.

Problem Statement

- In only few approaches, the interaction with cartographic interfaces has utilized gaze as an input (Giannopoulos et al., 2015; Giannopoulos, Kiefer, & Raubal, 2013).
- … research is missing on how alternative human-computer inputs like hand gestures in a Mixed Reality (MR) environment can be substituted by gaze-control for user-map interactions.

Research Objectives

MO TO DEVELOP
gaze-based interactions to facilitate user-map interaction in the MR environment

O1 TO IDENTIFY
user-map interactions for the gaze control in the MR

O2 TO ASSEMBLE
MR interfaces for the selected interactions

O3 TO EVALUATE
the performance and user experience of assembled interfaces.

Research Questions

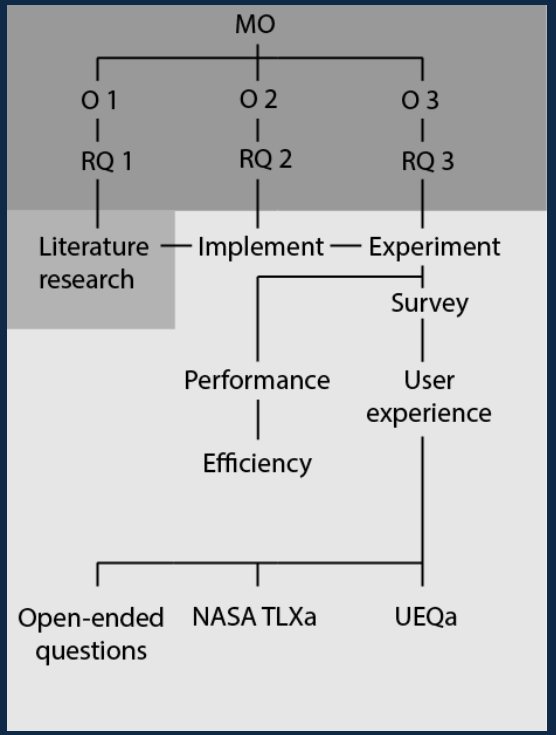
RQ1 How can users interact with the map in a Mixed Reality environment?
- a) What are the fundamental cartographic interactions?
- b) How can eyes control interactions?
- c) What are the fundamental interactions in the MR environment?

RQ2 How are MR interfaces assembled for the selected gaze-based interactions?
- a) What are the limitations of the developed gaze-based interfaces?
- b) What are the challenges in the development of the selected gaze-based interfaces?

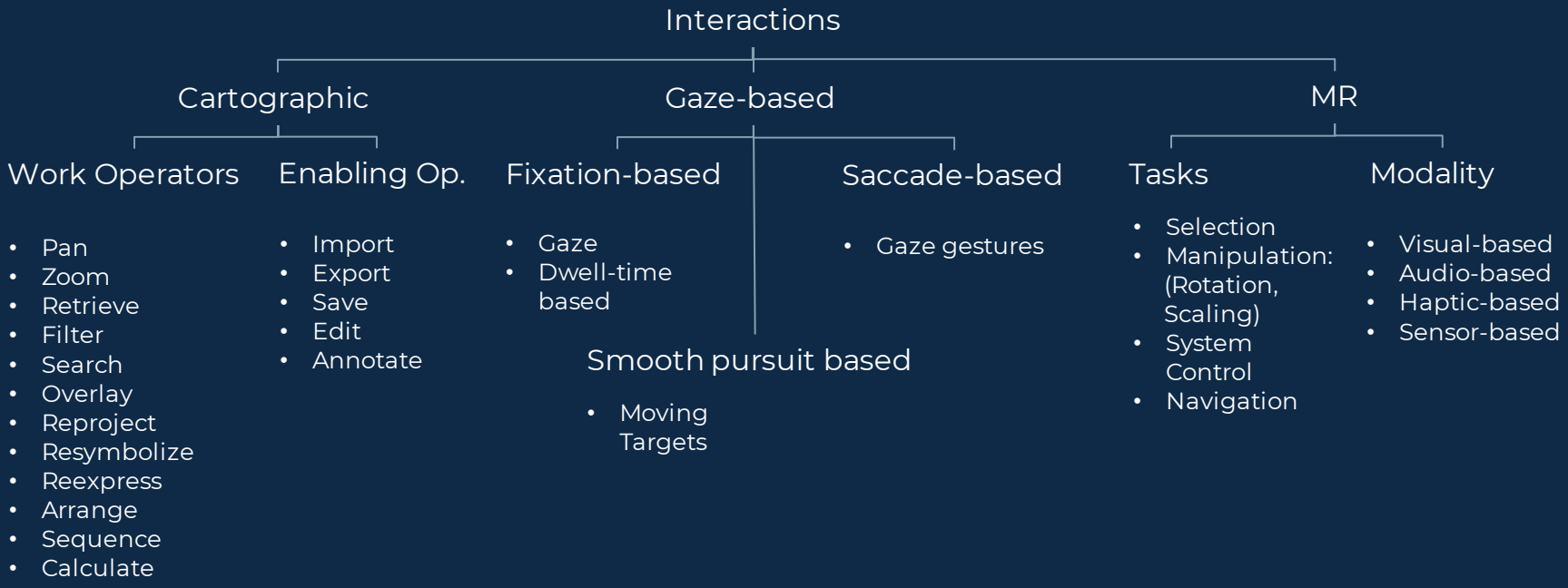
RQ3 How effective are the implemented user-map interactions in MR?
- a) What is the performance of the assembled interfaces?
- b) What is the user experience with the implemented user-map interactions in the MR environment?

Research Approach

Structure


MO
O 1
O 2
O 3
RQ 1
RQ 2
RQ 3
Literature research
Implement
Experiment
Survey
Performance
User experience
Efficiency
Open-ended questions
NASA TLXa
UEQa

Literature Review – RQ1

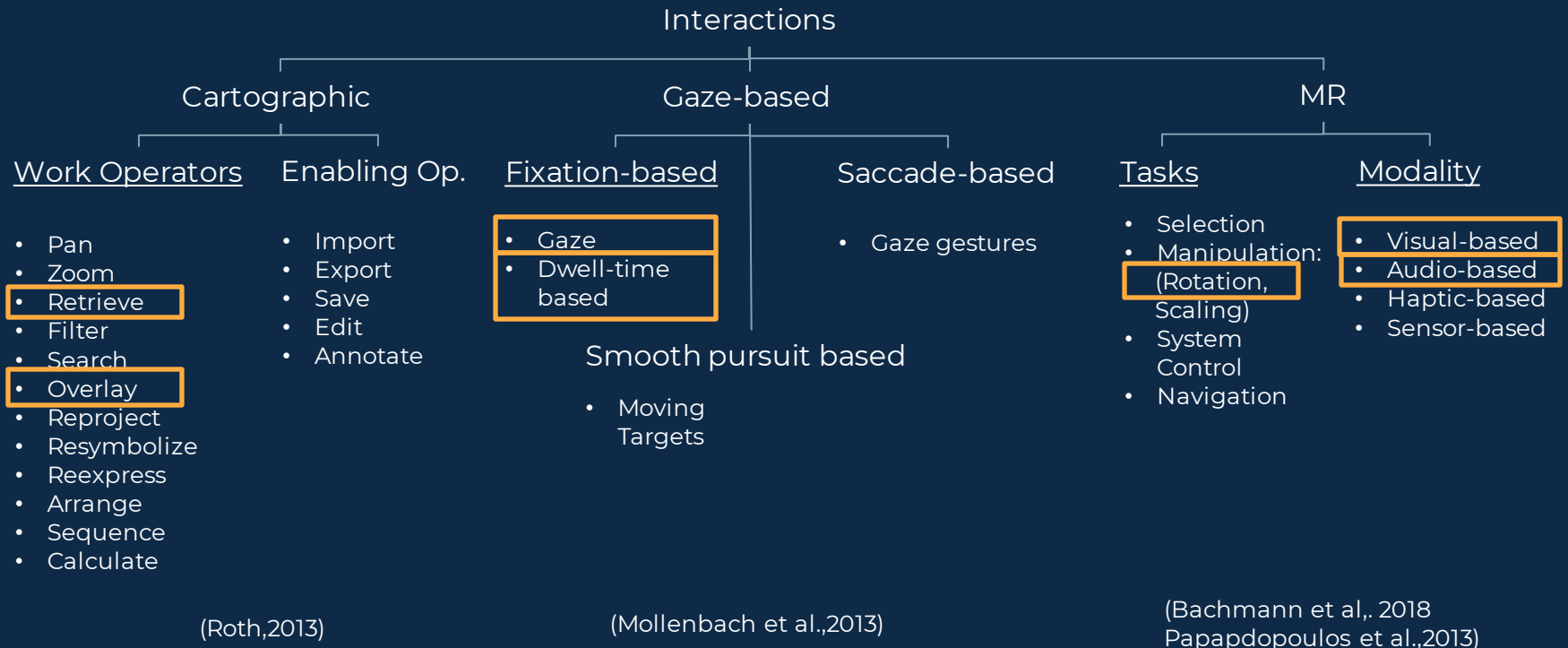


(Roth,2013)

(Mollenbach et al.,2013)

(Bachmann et al,. 2018
Papapdopoulos et al.,2013)

Literature Review – RQ1

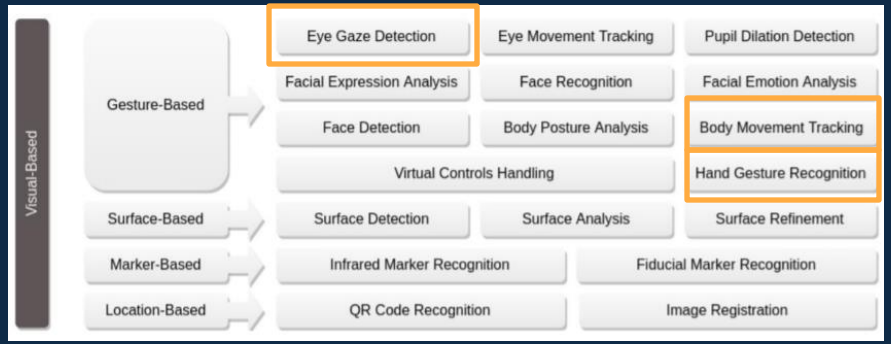


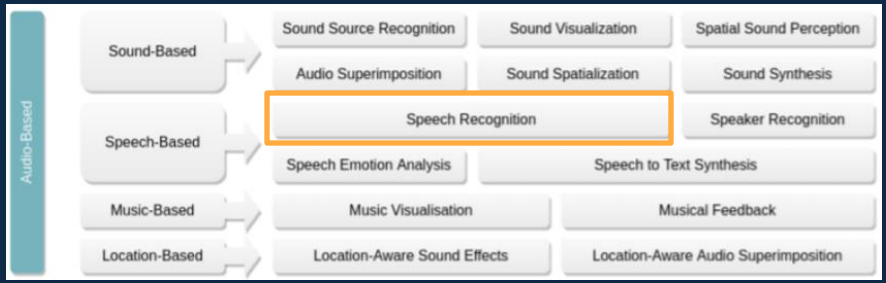
(Roth,2013)

(Mollenbach et al.,2013)

(Bachmann et al,. 2018
Papapdopoulos et al.,2013)

Modality – Context – Interaction (Papadopoulos et al., 2021)





Case Study

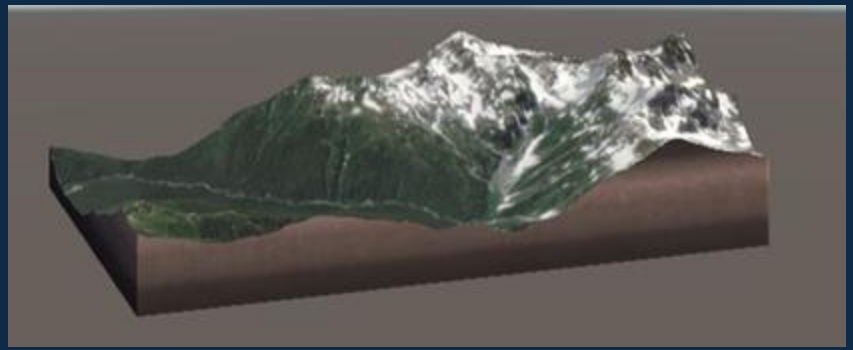
Implementation – RQ2 - Interfaces

	Conventional modality | action	Gaze-based modality | action	Gaze-aware modality | action
Overlay	 air-press	 dwell	 voice command
Retrieve	 point	 gaze	 gaze
Rotate	 point pinch move	 gaze	

Implementation – RQ2 – **3D Maps**

City

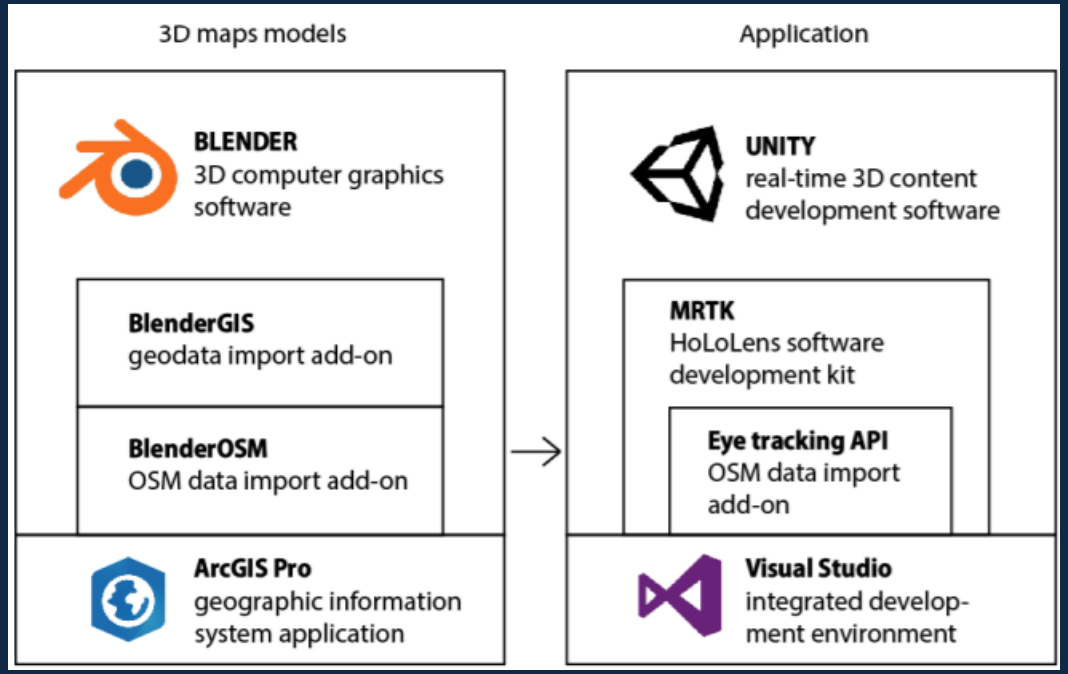
- Retrieve
- Overlay

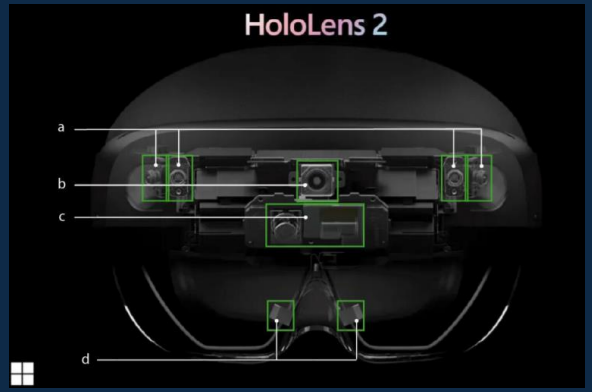


Terrain

- Rotate

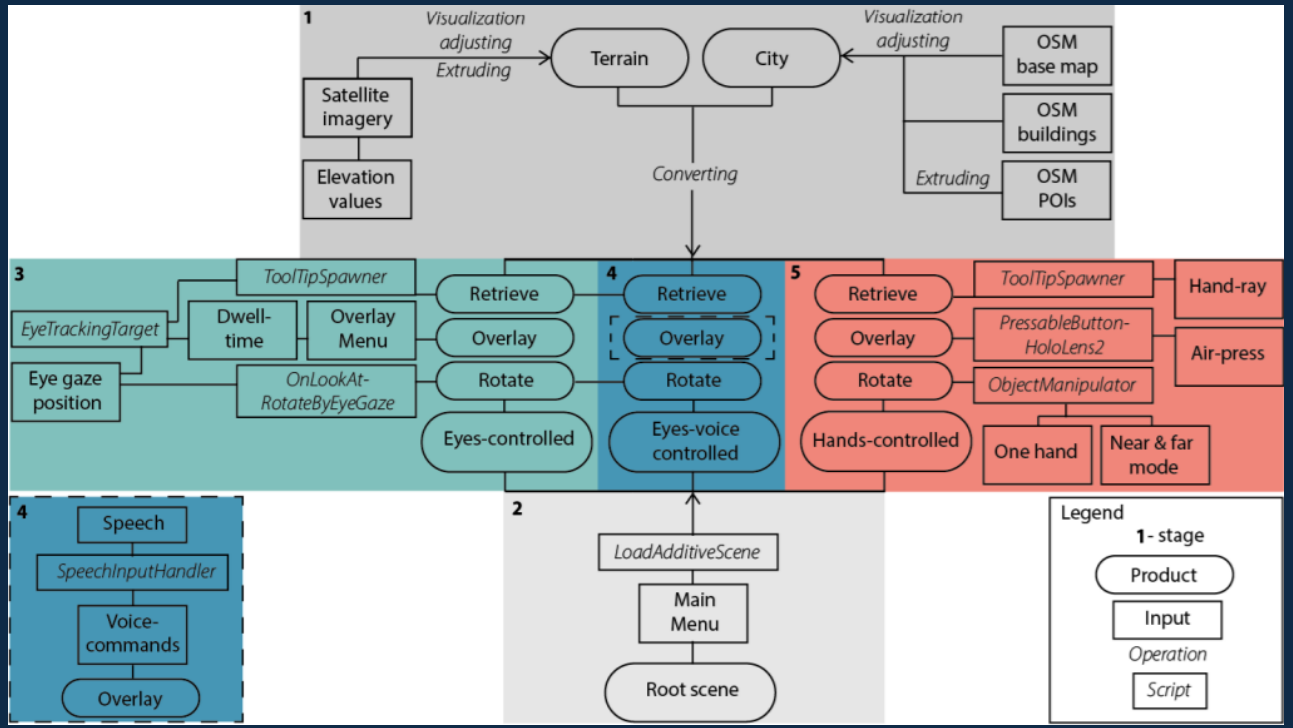
Implementation – RQ2 – Utilized Software & Hardware





HoloLens 2 cameras setup (front view)
a) 4 head tracking cameras - stereo and periphery;
b) RGB camera for photos/videos;
c) depth camera – near and far range;
d) 2 eye-tracking cameras – infrared.
Adapter from Microsoft

Implementation – RQ2 - **Workflow**



Experiment – RQ3

1. Pre-study questionnaire
2. Calibration
3. Interface 1 exploring
4. Tasks
    - Terrain
        - rotate the terrain to the task position
    - The city
        - turn on the satellite view and add the hotels
        - activate the name tag for the Sky Tower

    Recording the performance

5. User experience questionnaire

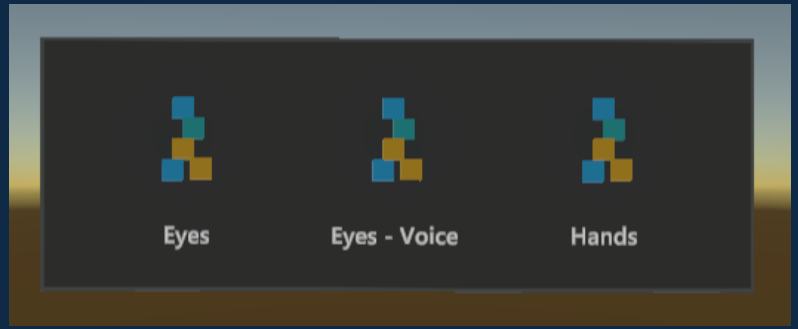
iterate for the 2nd and 3rd interfaces

Experiment – RQ3

Gaze-based Interface (eyes-controlled, eyes)

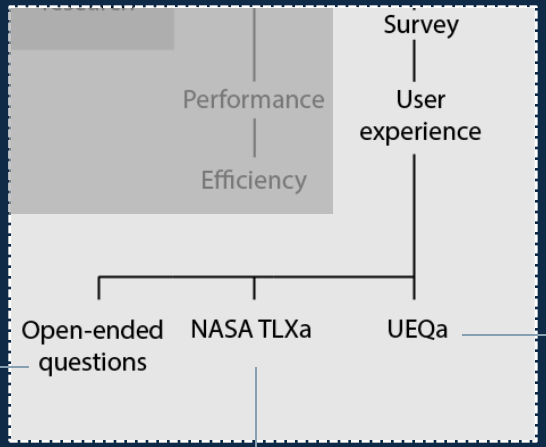
Gaze-aware Interface (eyes & voice -controlled, eyes-voice)

Conventional Interface (hands-controlled, hands)



Going to

Survey – RQ3



Difficulties
Suggestions

Mental demand
Physical demand
Effort
Frustration
Performance

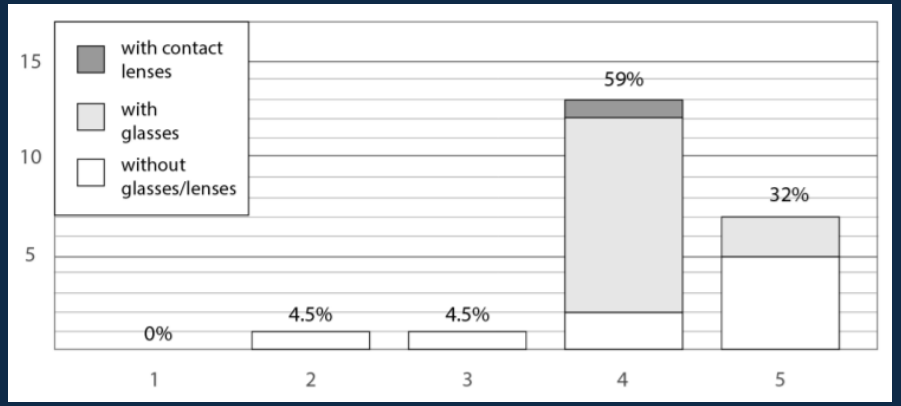
5 points ranking

annoying	enjoyable
Impractical	practical
boring	exciting
conventional	inventive
diff. to learn	easy to learn
confusing	clear

Results & Discussion

Participants

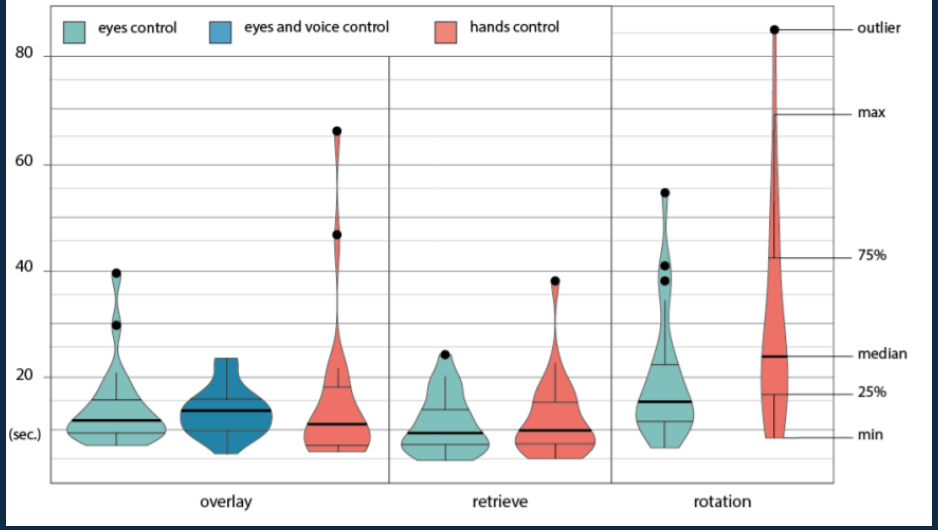
- 24 users
- 22 records
- 77% use maps often, 33% sometimes
- 54% experts or experienced with GIS
- 27% used MR devices once or several times
- 1 user uses MR device often



Performance

Table 6: Overlay, retrieve, and rotate tasks' completion time in seconds

	Gaze-based	Overlay	Retrieve	Rotate	Conventional	Overlay	Retrieve	Rotate	Gaze-aware	Overlay
Min		7,28	4,48	6,9		6,09	4,82	8,66		5,68
Max		39,33	24,32	54,62		66,25	38,12	85,66		23,56
Mean		14,27	11,33	20,02		15,43	12,38	30,34		13,68



- The outliers - the Midas touch.
- Individual approach to assigning the dwell time
- The gaze prediction area increasing
- Stopping the rotation
- The voice-command processing time.
- The voice command in a sentence or without pauses.
- Confusing press with tap
- Learning the rotation gesture

User experience – User Experience Questionnaire

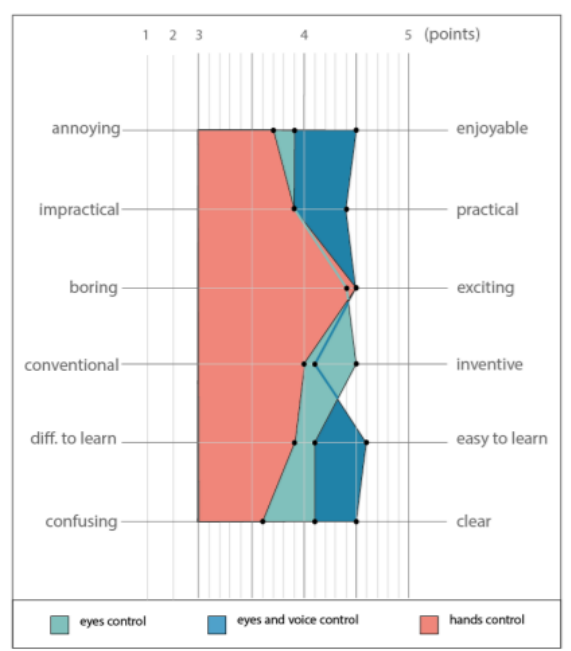
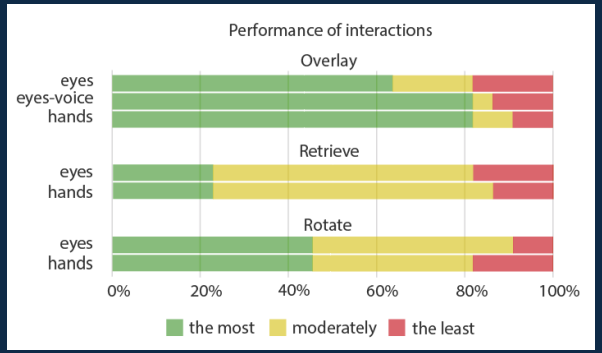
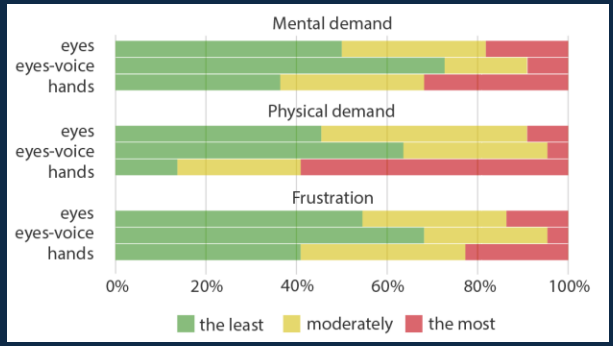


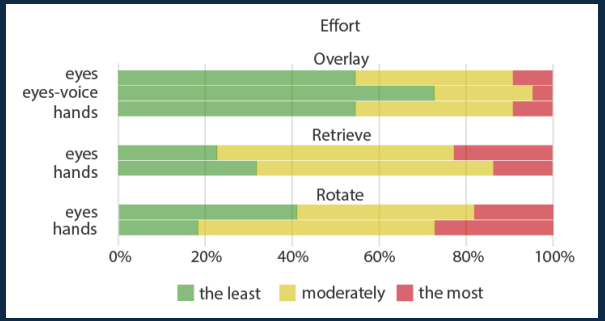
Figure 34: Visualisation of UEQa results' mean values

- Excitement to use interfaces (0.1)
- Activity preferences
- Challenging gesture learning
- Common technologies affect the inventiveness

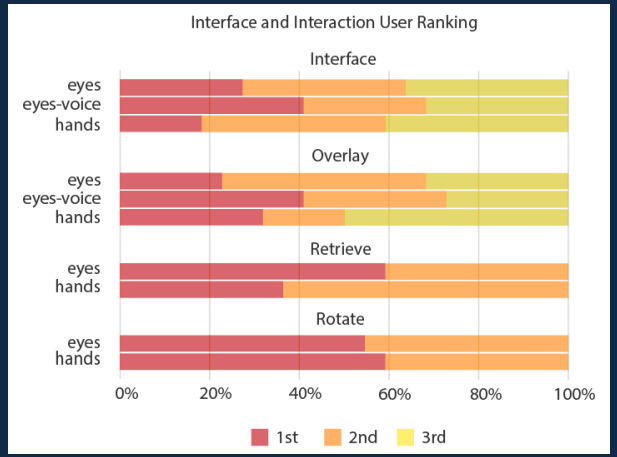
User experience – Task Load Questionnaire



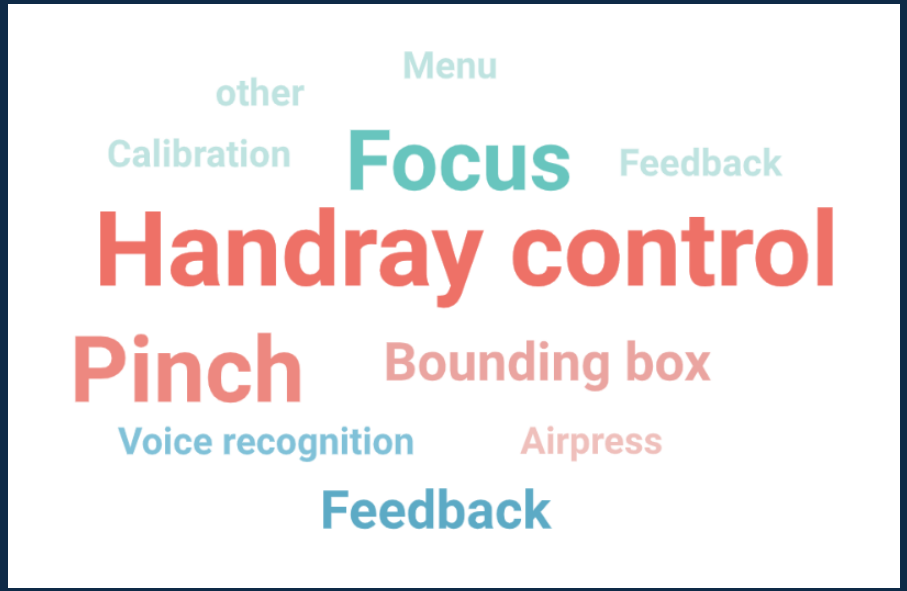




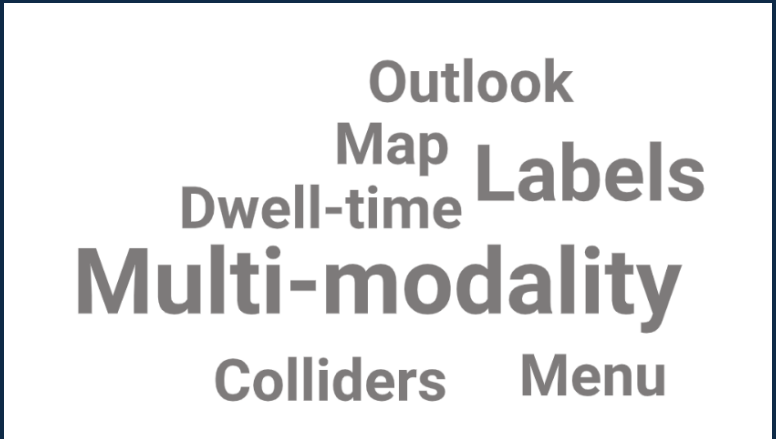
Interface Ranking



Difficulties



Suggestions



Challenges & Limitations

Combining the gaze and voice modality for the interactions

Constant order of the interfaces

This can present a limitation for the comparative evaluation of the three interfaces.

The visual and audial feedback design for the gaze-based and gaze-aware interfaces

The incomplete design of the maps: colliders, not restricted rotation.

This can influence the user experience results.

Conclusion

The fastest retrieve and overlay when using gaze as a controller.

The overlay interaction performance (gaze-based) is lower than with the voice-controlled.

The gaze-based interface - the most inventive interface, more enjoyable, easier to learn, and less confusing than the conventional.
Inferior to the gaze-aware interface in the same qualities.

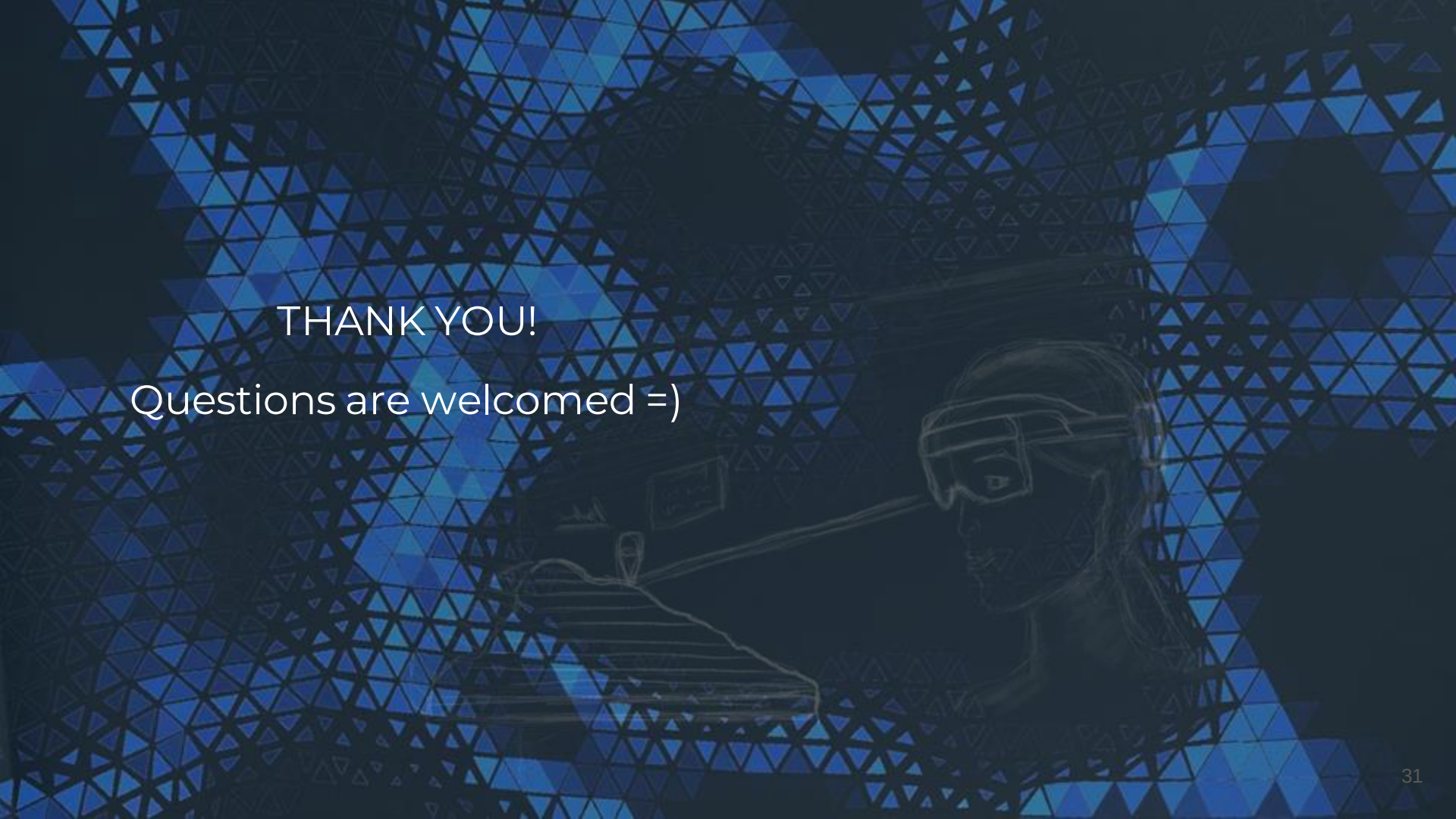
Requires less mental and physical demand and effort than the; however, but more than the gaze-aware interface.

The conventional - the most problematic interface. However, the gaze-based interface has also presented some difficulties, such as focusing on the target due to accuracy in the position of the gaze-pointer, or insufficient dwell time, or the incomplete map design.

Outlook

Improving the user experience of gaze-based user-map interactions

The multimodality can be explored for gaze-based interactions.


THANK YOU!
Questions are welcomed =)