LabelAR

Mixed Reality Project Proposal
Supervised by: Marcel Lancelle
October 4, 2024

GROUP MEMBERS

Matteo Boglioni | Tommaso Di Mario | Andrea Sgobbi | Gabriel Tavernini




I. DESCRIPTION OF THE PROJECT

The students will implement an app deployed on Magic Leap 2 (ML2). The app will display labels on landmarks placed over multiple locations of interest around Zürich. The app will be accessed from multiple locations that will be automatically recognized. The headset will self-detect the user's general site and display only the visible landmarks based on location and sensor data.

II. WORK PACKAGES AND TIMELINE

In this section, we'll present a detailed summary on how we decided to divide the project, allowing us to develop it progressively and in modular fashion. Each part will be developed and tested sequentially, allowing us to have, throughout the semester, a working stub of the final product while maintaining a cohesive structure.

1) **Label Placement and Short-Distance Occlusion.** To get acquainted with the headset and development environment, we will begin with the simple task of displaying labels within a small room. ML2 exposes an Occlusion API which makes use of local environment meshes to automatically handle occlusion of objects up to 7.5m away. For this simple task, we will avoid localization and just assume a fixed initial position.

2) **Localization.** Due to the absence of a GPS sensor on-board, we will need to perform image-based localization. A convenient option is using Google's ARCore API [2] which performs global localization using StreetView images. The problem with such an approach is the access to the API itself, which may be costly, but comes with the advantage of requiring no environment map. Due to potential compatibility issues of ARCore with ML2, a possible fallback with similar behavior could leverage GPS data from the user's phone: since the main use-case for our App is exploration, we consider access to a mobile phone a reasonable assumption.
The alternative to this is using a pre-scanned large-scale map of the area, such as GeoZurich [3]. Due to the large size of such a map, we would have to resort to a smaller local version, for example including the surroundings of the HG building. If the localization step proves too complex, we could also resort to a fixed location marker: these markers could be placed in various outdoor locations, and could also directly link to a local map as well as the relevant landmarks for the region. Otherwise, WiFi signal could be used to recognize the specific area and fetch the relevant map and landmarks.

3) **Long-Distance Occlusion.** Within a city environment, a lot of the distant landmarks such as mountains or smaller buildings won't be visible most of the time. To avoid visual clutter, labels should properly handle the occlusion caused by objects too far for the Occlusion API. The way this could be handled largely depends on how we choose to handle localization: if we use ARCore geospatial localization, we won't have an environment map which we could easily use to check the presence of occluding objects between the user and

the landmark location. In this case, we could resort to using the on-board cameras for stereo depth estimation using standard algorithms such as Block Matching and Semi-Global Matching or CNN-based approaches like PSMNet [1] and FadNet [4]. If we are able to obtain nearly real-time performance from such methods, we could project the landmarks onto the image planes to verify whether the captured depth is less than the real depth, indicating that the landmark is occluded.

4) **User Interface Improvements.** Once the main landmark functionality is implemented, we will begin working on some improvements to the User Interface (UI). To simplify exploration, we could display a small-scale 3D map with all the landmarks for the user to interact with; this map could either be the full pre-scanned map, or the local environment mesh built by ML2 (in this case, the landmarks may appear outside the mesh but still provide useful directional information). An occlusion toggle would enable the user to directly see distant landmarks in the real environment: using the mini-map, this could be activated for individual landmarks to avoid having all other landmarks showing up.

Including the time needed to adapt to the new environment, we estimate that the initial simplified task will take us 2 weeks. Although this initial task should only require working with Unity, the *Localization* task will most likely require integrating external code, as well as investigate the multiple options we have outlined. For this task we have thus allocated 4 weeks of work, leaving 3 weeks to work on the final details of *Long-Distance Occlusion* and *UI Improvements*. Andrea and Matteo will mostly be responsible for the initial setup, while Tommaso and Gabriel will focus on exploring the various localization ideas. Finally, the whole group will work together on the final touches to improve usability.

III. Outcomes and Demonstration

This project aims to develop a fully functioning app that will run on ML2. The final user will only have to wear the AR glasses and launch the app from the device home menu. Once launched, the app will locate the user in the map using its multiple sensors and cameras. Finally, LabelAR will start displaying the visible labels on the respective landmarks and update them based on the user movements in real-time. We plan to create a demo video from the multiple locations that will be enabled. Furthermore, we propose to give a real-time demo by placing one of the team members at the closest location and live-streaming the experience in first person.

References

[1] Jia-Ren Chang and Yong-Sheng Chen. Pyramid stereo matching network, 2018.
[2] Google Developers. Geospatial api with unity ar foundation, 2024. Accessed: 2024-10-03.
[3] Hayko Riemenschneider. Geozurich - computer vision laboratory, 2018. Accessed: 2024-10-03.
[4] Qiang Wang, Shaohuai Shi, Shizhen Zheng, Kaiyong Zhao, and Xiaowen Chu. Fadnet: A fast and accurate network for disparity estimation, 2020.