

Design and Implementation of Asynchronous Remote Support

Irene Reisner-Kollmann*

Andrea Aschauer†

University of Applied Sciences Upper Austria



Figure 1: Screenshot of the mobile app, where the annotations are directly overlaid in instant mode.

ABSTRACT

We present an AR-based system that allows spatial collaboration for multiple users. The main contribution is that on-site and remote users can work together in an asynchronous way. On the one hand, mobile users can create a new project by documenting the environment on-site and saving this data for asynchronous processing. On the other hand they can access previously saved data on-site with an augmented reality view where annotations and other content are directly visualized on top of the real world. Desktop users can access the data remotely and have more sophisticated ways for editing. Both, mobile and desktop users can revisit, view and edit a collaboration project anytime. We show how the system can be setup with standard hardware and software modules and share details about the implementation.

Index Terms: Human-centered computing—Human computer interaction (HCI)—Interaction paradigms—Mixed / augmented reality; Human-centered computing—Collaborative and social computing—Collaborative and social computing systems and tools—Asynchronous editors;

1 INTRODUCTION

For many services in industrial environments it is necessary to access a remote site multiple times and consolidate information from different people. For example, consider the following workflow for the maintenance of a large machine at a customer site. First, a customer service representative gathers all information on-site including pictures, videos and notes from the customer. Based on this information, employees in the back office determine what exactly will be repaired and create an offer for the customer. Additionally, they add more information like manuals or specific instructions. Finally, a technician executes the maintenance and retrieves all information from the project. If the machine needs maintenance again at a later time, it is beneficial to access the information again.

*e-mail: irene.reisner-kollmann@fh-steyr.com

†e-mail: andrea.aschauer@fh-hagenberg.com

We provide a system, where multiple users can access and edit spatial data at different times. We provide different interfaces for mobile users on-site and desktop users at a remote location. The system is mainly based on annotations on 2D snapshots (see Fig. 1), but other content types are supported as well. For mobile users, the content is spatially registered to 3D landmarks. For desktop users the content is attached to 2D images, as they do not have the full spatial information.

We identified the following goals for our asynchronous remote support system:

- Exchange information between local and remote locations
- Create a documentation and exchange information over time
- Include existing data from other sources in order to be easily accessed on-site
- Keep the mobile app simple by moving time-consuming tasks to the desktop app

The main benefit of asynchronous collaboration is that not all participants have to attend simultaneously. The participants can conduct their work at their preferred time slot. They have the possibility to carefully plan and review their input. Another practical problem, which demands asynchronous methods, is the poor internet connection in some peripheral industrial sites.

In this work, we leverage standard hardware devices and software components for creating a sophisticated collaboration system. This allows to quickly get advantage of the system as it does not require any costly investments or disruptive organizational changes.

2 RELATED WORK

Remote collaboration using a live video stream and augmented reality has been an active research topic in recent years [1] and has also led to commercial applications [8,9]. The systems usually connect a user with a mobile device and a remote expert. We use a very similar concept and transfer it to an asynchronous paradigm.

There are different variations of remote support systems [3,6,7]. Annotations can be stabilized in world coordinates or fixed to the video frame. They can be persistent or vanish after a short time. The expert's view is freely selectable or it is bound to the current view of the local user.

In one of the first papers, Gauglitz et al. [4] present a remote collaboration system for mobile devices. They coarsely reconstruct the 3D scene in the remote application for correctly placing 3D annotations. The remote user can select any of the previous view points provided by the local user.

A comprehensive review of collaboration techniques in augmented and virtual reality is given by Ens et al. [2]. They analyze mixed reality systems according to concepts of Computer Supported Cooperative Work (CSCW) research. The vast majority of the analyzed papers, namely 95 %, focus on synchronous collaboration.

Irlitti et al. [5] describe theoretical considerations and challenges for asynchronous collaboration using augmented reality. They suggest the importance of comprehensive information from users and for the spatial and temporal organization of the data. In contrast to our work, they mainly focus on local collaboration and not on remote users.

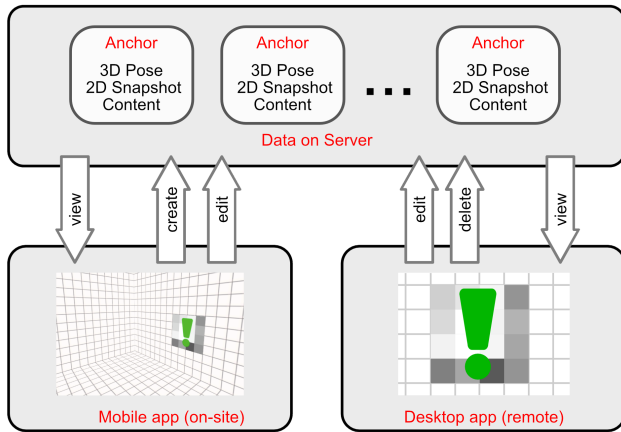


Figure 2: Overview of collaboration system: Mobile and desktop users share the same content on the server.

3 DESIGN CONSIDERATIONS

Fig. 2 shows the structure of our asynchronous collaboration system. All users access the same data on the server. All data is organized in anchors, which contain a 3D pose, a 2D snapshot and the content itself. Only mobile users can create new anchors, because the 3D information from the augmented reality framework is needed. Deleting anchors is only possible for desktop users, because we want to keep the mobile app as simple as possible.

3.1 Content types

The most important content are 2D annotations. Both, mobile and desktop users can edit the annotations with a standard drawing tool on top of a 2D snapshot of the scene. Annotations are stored as separated images with a transparent background. Annotations allow to highlight specific objects in the scene or to scribble additional instructions.

A special form of the 2D annotations is text. The visualization is similar to other 2D drawings, but it is important to store it as plain text for editing and searching.

Audio content allows the mobile user to easily and quickly capture information. Video is useful to provide an overview of the whole scene or to include external data such as instructional videos. Documents allow to attach further information, e.g. the user guide for a specific machine. Documents are basically stored as links to external files, i.e. they can contain almost any type of data. The only prerequisite is that a viewer is installed on all devices.

The system can be easily extended to any other content type. We'd like to incorporate interactive elements in the future such as animated instructions or checklists.

3.2 Mobile app

The mobile app is used for creating, viewing and editing content on-site. We currently support handheld devices, but other AR displays such as head-mounted displays could be used as well. Fig. 3 shows the main user interface of the mobile app.

We use the standard augmented reality view, i.e. the user sees the live video of the device's camera together with the AR annotations. The information from the AR framework such as points and planes can be optionally visualized. This gives the user a better understanding whether the scene has been captured well enough. If it disturbs the user from the actual information, it can be simply hidden.

As noted before, our main content are 2D annotations. There are two modes for visualizing them in the mobile AR view:

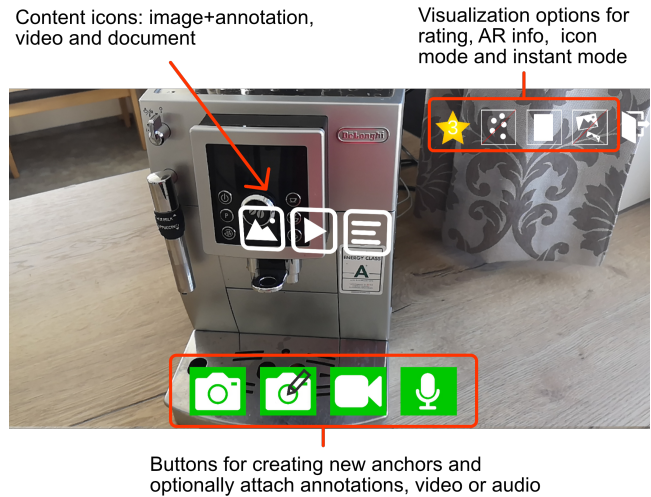


Figure 3: Screenshot of the mobile app in icon mode.

Instant mode As shown in Fig. 1, the 2D annotations are directly overlaid to the live camera image by rendering them as quads at the anchor's pose in the three-dimensional space. Multiple annotations are shown at the same time.

Icon mode An icon is shown at the 3D location of the anchor. When clicking the icon, the original 2D snapshot and the annotations are shown full-screen. This mode is especially useful in case of poor registration (see Sect. 4.1), because direct overlays are not accurate enough. It also allows the user to see whether something has changed in the scene compared to a previous session.

Other content types are always indicated with an icon in the 3D scene and are opened by clicking on the icon. We use standard software for showing videos and documents.

The user can create new anchors by tapping one of the buttons in the bottom (see Fig. 3) or directly onto the video stream. The 3D position of the anchor is computed by raycasting the reconstructed 3D points and planes from the augmented reality framework. Optionally, the user can create and edit 2D annotations to the current snapshot with a simple drawing tool. Besides drawing lines with the finger tip, the user can select a color and clear the whole screen.

While our system is designed for connecting on-site and remote users, it is not mandatory to have a remote user. The mobile app can be used on its own and allows users to view and edit the data multiple times on-site.

3.3 Desktop app

The desktop app is aimed for remote users. They can review the data from the object or environment to work on, which has been provided by the mobile users. Further they can edit the data and provide additional information. Fig. 4 shows the main parts of the user interface for remote users.

The desktop app visualizes the data only in 2D. Therefore, the 2D snapshots are the main point of interaction. The user can flip through all 2D snapshots. Annotations are always overlaid if they are present. Other content types can be opened with buttons.

Users can edit the 2D annotations with a similar drawing tool as in the mobile app. They can add any number of additional content and assign it to one of the existing anchors. It is not possible to create new anchors in the desktop app, because the 3D information is not available.

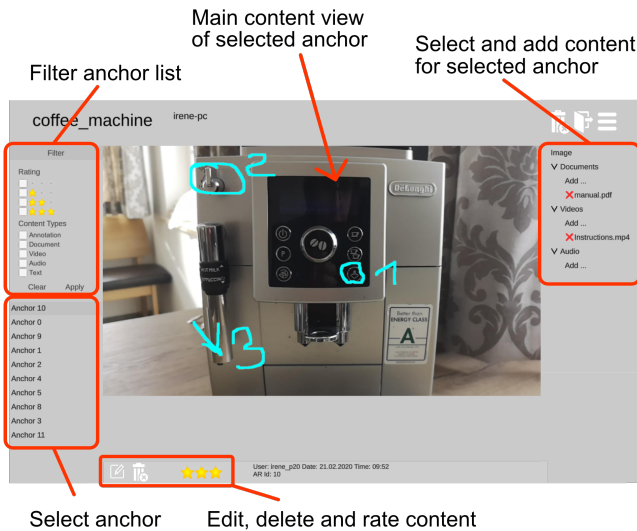


Figure 4: Screenshot of the desktop app.

Usually, the goal on-site is to capture as much information as possible because it's often difficult to access it again. The desktop app offers various ways for organizing large amounts of data. There is a rating mechanism in order to optionally hide less important anchors. The remote user has also the possibility to delete specific content or complete anchors.

4 IMPLEMENTATION DETAILS

The system is implemented in Unity 3D, which has the benefit that many parts can be shared between the mobile and the desktop app. We use Unity's integrated augmented reality framework ARFoundation. ARFoundation provides a unified API for ARCore and ARKit and thus supports both, Android and iOS. ARFoundation is also necessary for using Microsoft's Azure spatial anchors in Unity.

4.1 Registration

An important factor for the AR application is to register the current session to the existing spatial data. We provide three different approaches:

Cloud anchors Azure spatial anchors are created by capturing the surrounding scene and can be shared with other Android, iOS and HoloLens devices. When the scene is viewed again, the pose of the anchor is re-established without user interaction. This approach fails if there were too many changes in the scene. In our system, we use one Azure spatial anchor per project.

Marker Aligning the scene to a known marker is supported by all established AR frameworks. Usually, it is not possible to permanently mount markers at customer's site. In this case, the marker location can be inferred from screenshots of previous sessions. The same printed marker can be used for each support case and the user positions it again in the scene according to the screenshot.

Manual registration The user selects a screenshot from a previous session and manually overlays the current camera view. This procedure is clearly very inaccurate. We included it as a fallback when there is no internet connection and markers have not been used in previous sessions.

None of these registration methods is absolutely free from defects. If the registration is not good enough, the annotations will be overlaid

on wrong objects. In this case it is recommended to fallback to the icon mode and show annotations only on top of the original snapshot.

4.2 Data storage

An important request from industrial enterprises is that data can be easily shared with other software units. Therefore we decided to store all content in conventional data types. For example, all 2D snapshots and annotations are stored as individual image files. The same is done for audio, video and documents. Additional data from our system is stored in json-files. This approach makes it easy to create importers for other software systems or to view the data directly with off-the-shelf software.

5 CONCLUSIONS AND FUTURE WORK

We showed a system that allows on-site and remote users to exchange information on a specific site. Mobile users can see the information directly overlaid to the real object of interest in an augmented reality view. Remote users have a 2D view on the data and have more ways for adding additional data. We showed that the system can be simply implemented with standard hardware. We think that asynchronous remote collaboration has a big potential of improving collaboration on visual data over distance and time.

In the future we plan a detailed evaluation of the system. We like to compare it to traditional data sharing and to synchronous remote support systems.

An obvious extension will be the support of head mounted displays. Another big improvement will be to make 3D data accessible to the remote user. The remote user might get a better overview and has the possibility to create new anchors. An important topic will also be large scene changes between sessions and the handling of incorrect scene registration.

ACKNOWLEDGMENTS

This work was funded by the FFG under the project Mixed Reality Based Collaboration 4 Industry (Collective Research).

REFERENCES

- [1] R. A. J. de Belen, H. Nguyen, D. Filonik, D. Del Favero, and T. Bednarz. A systematic review of the current state of collaborative mixed reality technologies: 2013–2018. *AIMS Electronics and Electrical Engineering*, 3(2):181–223, 2019.
- [2] B. Ens, J. Lanir, A. Tang, S. Bateman, G. Lee, T. Piumsombon, and M. Billinghurst. Revisiting collaboration through mixed reality: The evolution of groupware. *International Journal of Human-Computer Studies*, 131:81–98, 2019.
- [3] O. Fakourfar, K. Ta, R. Tang, S. Bateman, and A. Tang. Stabilized annotations for mobile remote assistance. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pp. 1548–1560, 2016.
- [4] S. Gauglitz, C. Lee, M. Turk, and T. Höllerer. Integrating the physical environment into mobile remote collaboration. In *Proceedings of the 14th international conference on Human-computer interaction with mobile devices and services*, pp. 241–250, 2012.
- [5] A. Irlitti, R. T. Smith, S. Von Itzstein, M. Billinghurst, and B. H. Thomas. Challenges for asynchronous collaboration in augmented reality. In *2016 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct)*, pp. 31–35. IEEE, 2016.
- [6] A. Jakl, L. Schöffner, M. Husinsky, and M. Wagner. Augmented reality for industry 4.0: Architecture and user experience. In *FMT*, pp. 38–42, 2018.
- [7] S. Kim, M. Billinghurst, and G. Lee. The effect of collaboration styles and view independence on video-mediated remote collaboration. *Computer Supported Cooperative Work (CSCW)*, 27(3-6):569–607, 2008.
- [8] Microsoft. *Remote Assist*. <https://dynamics.microsoft.com/de-de/mixed-reality/remote-assist/>.
- [9] PTC. *Vuforia Chalk*. <https://chalk.vuforia.com/>.