TeamAR – Generic Interface for Cooperation Using Augmented Reality

Dawid Pacholczyk
Department of Software Engineering
Polish-Japanese Academy of IT
Warsaw, Poland
dpacholczyk@pjwstk.edu.pl

Mariusz Trzaska
Department of Software Engineering
Polish-Japanese Academy of IT
Warsaw, Poland
mtrzaska@pjatk.edu.pl

Abstract—Thanks to technology, the world is getting smaller and smaller every day. We are moving towards the model in which we do not have to be in the same place where our employer is located. Remote, scattered teams are something that we encounter every day. In our research, we explore methods and strategies that will allow to increase the efficiency of teamwork using augmented reality. We are looking for a way to bring a low-cost software platform that will allow to create augmented reality sessions and bring 3D objects into the world of each person. These objects should be shared but also be independent on some level for every participant. Such platform can help with the problem of availability of augmented reality for a wider audience. This paper provides an overview of research done in this field and our thoughts related to the topic and description of a working prototype.

Keywords-Augmented reality; Collaborative augmented reality; Cooperation.

I. INTRODUCTION

Augmented Reality (AR) allows users to see the real world together with some generated, virtual objects. Using this technology, we are capable to present not only simple 2D overlays, but also complex 3D structures and animations. There is an old saying: “A picture is worth a thousand words”. If we can present complex structures and animation as a part of our world, why not use them to improve our communication and teamwork?

In our research, we are looking for ways to improve remote cooperation and resolve the problem of low availability of AR. Our goal is to create a generic platform that will allow to start a quick AR session without any programming or need for special equipment. This project is intended for companies, their teams and average users.

In our opinion, it is much easier to just present a complex object and experience it through some interactions rather than trying to explain it. For each participant, we want to provide their own perspective in addition to the shared state of the scene. During the presentation, each person can reach his/her own goal without disturbing the rest of the team. We want to maximize the efficiency of the team and minimize the technology adoption cost.

Such platform can have a positive influence on elements such as: lowering the cost of creating complex mockups, or new kind of low cost user experience tests. In the design sector, it helps with easier presentation as well as keeping low costs. Positive influences of the platform are also envisioned for the education and entertainment sectors.

The rest of the paper is structured as follows. Section 2 covers the background. We describe how ideas related with cooperation using AR evolved in time. In Section 3, we describe the concept that stands behind TeamAR, what are the goals that we want to achieve. Section 4 describes the implementation details and architecture that was developed for TeamAR prototype. As this is the first prototype of our platform, in Section 5 we describe the future work that stands in front of us. Section 6 is the conclusion, the summary of what we already achieved and what possibilities are created thanks to our system.

II. BACKGROUND

Augmented reality is a very broad field for research. Over the past few years, several studies explored the usage of augmented reality in our everyday actions, starting from simple usage like games [16][17], entertainment [8], productivity [29][30], through more complex tasks like data visualization [33], repairing [31] to logistic support [32].

Our system was inspired by several previous research projects in augmented reality and possibilities to use them for collaboration and teamwork.

Shared space [1] presented how AR can become a powerful tool in face-to-face meetings. It was one of the first significant steps in this field. Shared space allows a group of people to work in one room on common 3D objects. Users must be equipped with a HMD (Head Mounted Display). The workflow, and objects attached to a 2D marker are predefined. Each user sees the same object as every other in the room, they can share, and move the markers in the same way as they would do with normal elements of the environment. Shared space showed how much AR can help us to enrich our communication, and methods of work.

Our project took the lead in finding the solution for removing one of the biggest obstacles, which is lack of flexibility, and strong need of highly qualified staff. We also want to provide a possibility to work remotely if needed.

Project Studierstube [14] increased mobility of the AR systems. Authors built a mobile workstation, combined with optical see-through HMD. Thanks to that, multiple users can cooperate with each other, working on the same 3D objects with shared state. Each user has its own version of the object, set in an individual coordinate system. Thanks to that, one user will not affect the visible state of the second user.
Although this project was built using a laptop, modern mobile hardware is powerful enough to reduce the weight and size of AR devices. From this work, we can specify five key advantages, and elements that need to be met in similar platforms. Those elements are: virtuality, augmentation, cooperation, independence, and individuality.

Cooperative AR is not only helping in communication itself, it can be much more than that. For example, *Transvision* [12] is focused on supporting designing tasks. It is based on computer generated images displayed on palmtop (video see-through strategy). The augmentation helps designers as it combines two important elements: the possibility of an easy and cheap design change and physical contact between participants.

Similar task was taken by the team of the *ARVita* [13] project. It shows how AR can help engineers in their everyday work, by adding new powerful tools which they can use. Our work methods evolved from simple notes, and plans, through photos, up to 3D/4D Computer Aided Design (CAD) models. *ARVita* is trying to go one step further, namely, it adds a 3D animation to the physical world to give us a better perspective on the work that needs to be done. This idea allows to prepare complicated simulations and to watch them in full 3D from different perspectives.

Many researchers are aware of the benefits, and possibilities which AR can deliver in terms of cooperation whether face to face or remote. Augmented reality can even improve our chances to save lives in the face of crisis. In [3], there is an example how AR can be used to coordinate work of different organizations, and how it can help to make the best decision based on very dynamic circumstances. Despite the fact that the described scenario is not very flexible, and it is hard to reproduce in a short period of time, the results of this research are very promising. It is worth noting that such a system can be enriched with remote collaboration. We believe that decreasing the access time to specialists can help people particularly in crisis situations. Unfortunately, this solution is faces similar problems to those previously mentioned: lack of flexibility, time consuming tasks needed for preparation of the scenario, complicated hardware. In our project we are trying to remove those complications.

The natural path of the evolution of collaboration using AR is to allow a remote user to access the same state of the object. We can see such first attempts in [2], where the authors implement a simple Tetris game that allows more than one user. This research makes us aware how hard it is to achieve such goal. It reveals key problems such as sharing the object state, anchoring it in a fixed place, and performing some interaction. Those tasks, as a single case are quite easy to resolve, the true challenge is to combine everything into one platform. Similar to other research, this one uses markers for setting objects in the scene and uses the same technique to support the interaction.

*MARS* [11] presents a path which shows how we can work in teams on separate tasks to achieve a common goal. This project allows to send extra information from people in the office (indoor) to a person in the field (outdoor). Everything is displayed on HMD (see-through). Such a concept provides multiple possibilities from engineer support during assembly tasks, providing help for the soldiers on the battlefield. Similar approach, but from a different perspective, was presented in the project called *God-like interaction* [27]. It also focused on putting users in two groups:

- Indoor users that have access to a tabletop projector display system,
- Outdoor users with Tinmith mobile augmented reality system [28].

In the case of these studies, it is interesting that we can put different real objects in the perspective of the outdoor user, by capturing them in a series of photos focused on the table surface. Those objects are sent to the outdoor user and reconstructed on his/her display. Such an approach allows us to put virtual signposts, alerts, extra information, etc. We can also inform about important places to visit or areas that should be avoided.

Both *MARS* and *God-like interaction* draw attention to the fact that such systems have enormous potential to help during different crisis situations. We can treat indoor users as crisis staff (see also [3]), and the outdoor users as the rescue brigade. Thanks to augmented reality we are able to send to the people in the field much more information than only voice messages. We can mark where they should go, where they can find something important, etc.

All presented research projects have one thing in common, namely interaction. Regardless of what kind of project we are building, what hardware we will use, we always need a way to interact with objects that augment our reality. This is the very basic concept of AR. Nowadays, we have access to multiple tools which allow us to work in teams irrespective of location (Google Docs [22], Office 365 [23], etc.). We need to remember that working on different computers, even in the same place and on simple tasks, can create problems. Lack of the same perspective, and a common view, limits ways of our communication, e.g., we have no option for using simple gestures or pointing at objects.

Another thing common to the presented projects, is a method of setting virtual elements in the real world. To make AR as much natural as possible, we need to achieve full transparency. Objects should behave in a predictable way, with a fixed position “glued” to the part of the scene. In many of presented papers, simple 2D markers were used to handle this as a well-known strategy. One of the disadvantages of this pairing the marker with an object preparing it, before we can start the work. In [15], we can see an attempt to do it in a more generic way. A shared vision system is a platform that allows to use dynamic markers made from the first frame from camera view and track it to display a 3D model on it. The whole thing is shared between thanks to a database in the cloud.

We cannot forget about the latest products of Apple, and Google: ARKit [36], and ARCore [37]. Using standard smartphones, we are able to prepare a simple map of planes in our environment, remember it, and place objects on them. Augmentation is very natural, and the results give a new hope in terms of popularization of AR. This solution provides the full virtualization of our perspective, virtual...
objects are placed on a virtual plane, we can interact with them physically.

Last but not least, let us remember that one of the best reference is how the market accepts a technology. It is hard to push such a complicated technology to people that are not related to science or any research. That is why a simple mobile game is the best scenario. Online, multiplayer game is a fantastic example of a remote collaboration between users in different locations. For example, Ingress [16], or Pokemon GO [17]. Both games are based on a simple concept, but putting them all together in a single application, available for over 2 billion devices [18], give us easy to use augmented reality solution for everyone. Both games convert the real world into a playground for every player. Their foundation is a collaboration between multiple users in the real time. They interact with each other physically, and they can interact with the same virtual objects on map. Massive popularity of those games shows how big impact augmented reality can have on our lives. Those simple games are a small proof of the fact that people can and want to collaborate using new technologies, and that using proper approach to the scenario can give a fantastic result.

III. TEAMAR - CONCEPT

In this section, we will describe concept of our system TeamAR. We will describe the ideas, goals, and things that we want to achieve regardless to the current state.

A. Concept

In this section we will describe concept of our system TAR (Team Augmented Reality). We will describe the ideas, goals, and things that we want to achieve regardless to the current state.

Our project is focused on achieving the following goals:

- First, and most important. TeamAR is a project that must be usable in real life scenarios. We are focused on preparing usable prototype that can be easily implemented in every company, and different environments, that can get value from it. Every decision made must be compatible with this requirement.

- The system must be easy to use, and flexible. It must allow multiple users to work on a shared task.

- No extra programming needed. User only configures a session with selected markers, and objects. All participants join the defined session using its identifier. The whole process is supported by our platform.

- No extra hardware, except a smartphone, is needed. Using a head-mounted display (HMD) or special sticks with marker for manipulating is unnatural, and may cause problems with configuration, or may act as a deterrent for new users. We also want to avoid additional devices, because currently available ones are expensive, uncomfortable, and hard to get. That is why our goal was to create a platform that will use smartphones only.

- Ready “on-fly”. TeamAR must be easy to use and easy to manage. That means that application must be able to learn new patterns and objects during the runtime. No recompilation and even no restart of application should be necessary.

- The whole system must be mobile. We do not want the user to be limited to just one place and surrounded by cables.

B. Platform

TeamAR is based on the SaaS model. We built a web application for managing session and user synchronization. We used a standard smartphone with Android OS for working with the sessions (augmentation). Thanks to that, we could solve four major problems:

1) User perspective

- **Access to hardware** – nowadays, almost everybody has a smartphone. We want to make our project as flexible, and easy to access, as possible.

- **Interaction layer** – as we showed in the previous parts, there is no common, easy for user, and a hardware-free way to interact with the generated objects. Thanks to a screen of a smartphone, we get such a layer without any extra devices. Besides that, this platform is already well-known to users, so there are hardly any barriers to entry.

- **Progress** – software (Android), and hardware (smartphone) platforms will evolve in a natural way. Thanks to that we will obtain new capabilities without cost increases.

2) Software perspective

**Easy management** – thanks to choosing a SaaS platform, we have a platform that allows to create and edit sessions, as well manage them. The platform is scalable, accessible, efficient, and reliable.

C. Augmentation

TeamAR uses smartphones, built-in cameras, and sensors. The output is seen on the screen (video see-through strategy) after augmentation with virtual objects. To pin an object in space we use 2D markers. When the user looks at the card with a pattern, computer vision recognizes it, links with the 3D object from the session, and displays it on the card. The details of the current state of the implementation are described further in the paper along with more technical details of future development. We selected the 2D markers approach, because they are easy to use. They can be sent as a link, and displayed on the screen, or just printed. Every team can have their own set of markers, and just connect them with different objects in a new session or even update the reference in the current session.

D. Collaboration

TeamAR will allow user to refer to every physical element that he/she would normally use in a normal face-to-face meeting, like personal notes. Thanks to AR, he/she will also be able to refer to objects that normally would be impossible to use. This project will combine most important elements from the two worlds: virtual, and physical.
Furthermore, our project will allow teams to work remotely or stationary in the same room. No extra configuration will be needed. Thanks to the same objects base, and state sharing, they will see the same things independently from the location of each user.

E. Interaction

As we already mentioned, we decided to use a popular software, and hardware platform. It provides a well know, widely used interaction layer: the screen. The combination of mobility, good user experience, and an easy to use device provides a flexible tool to work with. The application will allow interaction with all virtual objects. This information can be synchronized between all participants of the session. As you can see the interaction is very similar to pointing at an element during a normal face to face meeting. This approach makes possible to avoid dedicated programming for every scenario.

IV. IMPLEMENTATION

In this section we want to describe our current state of the development.

As mentioned in the previous part of this paper, TeamAR is based on two main parts:

- Web application - responsible for creating/updating the session, sharing, and synchronizing the state of each object.
- Mobile application - marker recognition, displaying objects, and interaction. Speaking briefly, it is used for augmentation.

We prepared a tool (web application) that allows to create a session (Figure 1). Even a user without any technical skills can prepare a configuration of the meeting that will use AR technology. Users can upload a set of 3D objects and connect each of them with a specific 2D marker. Such configuration will be propagated to every user in the session.

Every person who possesses the identifier of the session can connect from the level of a mobile application and participate in the meetings. Thanks to our approach it is not important if that person is in the same room or in the other parts of the world. Everything he/she needs is a stable internet connection (mobile connection is enough) and 2D markers related with the session printed or displayed on a screen. Application learns new patterns during the runtime and displays the 3D objects on proper markers basing on the configuration that it received.

We decided to use 2D markers over the current solutions like ARKit [36] or ARCore [37], due to the fact that it is more natural. Secondly, ARKit and ARCore are available on a small number of costly smartphones, and, as we want to create as available as possible platform, it is very a important factor that we had to take into account. Both mentioned solutions are very interesting and have big potential, but they create more virtual environment. Virtual objects are mapped and placed on virtual planes. These methods limit ways of interaction between participants of meetings, and even between specific persons, and virtual objects, for example a person cannot lift the marker as he would lift a physical object to have a better look at it.

We used Firebase Cloud Messaging as part of our infrastructure for easy and real-time synchronization of actions performed on the objects. If the host changes color of the selected object, or if he/she marks an object it will be instantly synchronized to all devices that are currently in the session. Each participant will see the action in the same way as he/she would see when someone points at a physical object in the room.

A. Communication

We use a cloud service to implement fast, real-time, and reliable communication layer between the web application, and the mobile device. As a cloud service, we selected Firebase Cloud Messaging (FCM) [34]. The main reason for this selection is that, we plan to support other mobile platforms, and FCM is a cross-platform solution. The whole architecture and communication pattern is visualized on Fig. 2.

The web application is divided into two parts. The first one, contains a user interface. It is used to manage the sessions. It will be deployed, and available for users as service.

The second one is a REST API used for communication between mobile application and the database, and/or cloud. Thanks to the modular architecture, our system is easy to maintain, develop, and upgrade in the future. We also achieved a situation in which we are not tied up to a specific technology. As we have mentioned before, we are working on creating a proper concept of team collaboration with the use of augmented reality, not on specific technology. With better tools, better technology, we can replace single modules of our system, keeping all functionalities intact. We believe that minimizing dependency on current solutions, and technological trends is the key to our goal.
B. Data exchange

One of the most important decisions that we had to make, was selecting proper types of data used by the system for session configuration, both for the marker, and the 3D object. Selected type for each of these elements had to meet the specific requirement: a popular format already available on the market. To make the whole platform easy to implement in a company/team, we need to avoid situations when the selected data format is not known or hard to process; easy to use by a common user. Integrating users with the session should be as easy as possible to lower the entry barrier.

Because of the above requirements, we have focused on types that can be transformed and saved as a text file.

For the 3D object format, we selected Wavefront .obj file [35]. It is easy to use, easily accessible, it allows to create simple objects without any experience on the user side. Wavefront file is a text file that can be uploaded to our application and stored in the database. Reading this format is fast, and Android platform delivers native tools to work with it.

Another important fact is that such 3D models can be uploaded to software like Unity3D and used in different types of application. This corresponds perfectly with our main assumption about using TeamAR in a real-life scenario. Such objects can be prototyped, shared, and discussed before performing other time-consuming actions.

Furthermore, this format is very universal and easy to send. All this combined creates a perfect selection for our platform. As it is text data we can in easily update mobile application with the information about new objects. The format is light, therefor we do not need a high-quality internet connection, all calculations are performed on the device locally basing on the provided data. Wavefront format is universal, so when we will migrate out application to other operating systems, we will already have proper tools.

C. Augmentation

Our augmentation is based on video-see through HMD concept. In this case we deliberately ignored the HMD for the reasons mentioned earlier in the text. The application generates 3D objects based on the definition of the model received from the API. Actions performed on each object are shared, every change to the object is synchronized between users in session but the perspective is individual. Each user can independently observe the object without making any impact on the perspective of any other user.

Each object is located only on one 2D marker. The marker can be printed and put on the table or displayed on the screen (Figure 2). This ensures flexibility. Beside that a new object can be connected to the same marker and overwrite previous settings. Thanks to that, there is no need to prepare a big number of markers, users can easily update current state of the session and work comfortably having just a few or even one marker.

D. Interaction

In the early stage of the project we decided to implement a simplified model of interaction based on a toolbox. Thanks to that, during our tests new users had quick overview of actions that can be performed. This also made the whole interface very natural. There is a possibility that this will be a standard form of interaction for TeamAR. The decision will be based on collected feedback.

Currently, our toolbox allows to perform simple actions: highlighting selected objects. It is especially important when the host wants to focus the audience’s attention on a specific object in the session with many different structures. The user can also perform color changing action, which allows to make a fast grouping of objects basing on a selected color. The user is able to change the color of each object that is available in the session. This feature was implemented to simplify discussion about multiple structures.

Of course, the state of all objects can be reset to the default one using the reset action.

The whole “action framework” is easy to extend and we will experiment with more types of actions which will increase capabilities of TAR.

V. Future work

Our future work will be focused on extending the functionalities of the prototype and delivering it to users.

We plan to create interaction based on touching particular fragments of the object and confront it with the current solution. Although we do not limit ourselves only to one technology, as we are working more on the concept, and the philosophy of making Augmented Reality more accessible rather than on specific software type.

In the current version, the interaction is limited to basic operation like highlighting and changing the color of a specific object. Of course, the whole solution works “out of the box”. No extra programming from the user is needed.

We will use the history of performed actions to allow disconnected, and/or new users to start the session with the same state of the object as each of the users.

After that we plan to perform some usability tests comparing the performance of standard cooperation versus AR collaboration.

VI. Conclusion

Collaboration with the use of augmented reality is a very important area of technology. In our opinion, this is the direction that should be chosen to improve of remote work/teams in many professions. Various scenarios force
different approaches, but our goal is to create a platform which will give the most possibilities.

It is obvious that hardware will change in time, also the approach may be a bit different because of the evolution of the technology, but the concept in most cases will remain unchanged. Similar situation can be observed in many researches quoted in this paper. Even some of our solutions were inspired by research made almost 20 years ago. Of course, nowadays they can be greatly enhanced with modern technologies and more powerful hardware.

This paper has two main contributions. The first one is to present the current state of the AR solutions for teams. We try to identify weaknesses and strengths.

The second one is related with the following question: why should anybody want to build such a system? We believe that our prototype answers that question. Most of the systems that we have presented in the Background section is complicated, hardware dependent, connected with preprogramming concrete scenario of usage, created only for a specific purpose. We fully understand that TeamAR will not cover every possible scenario, but it hopefully provides a set of general purpose tools. It can help popularize augmented reality thanks to lowering the entry threshold.

We hope that vision of such interesting and fascinating area of research like collaboration in AR, will encourage other teams and researchers to search new fantastic ways of interaction between people. We are sure that such systems will shape our future.

REFERENCES


