Holobook: the story through Hololens

Arriu Simone, Serra Sergio, Spano Lucio Davide

University of Cagliari, Italy

Mathematics and Computer Science Department

Abstract: In this paper, we discuss the implementation of a mixed reality application developed for Microsoft Hololens using the Vuforia library. The aim of the project is to bring children closer to reading without neglecting the technological aspect while maintaining firmly the concept of not digitizing it. The focus is on the interaction between user and holograms through a guided learning on the different types of interaction (gesture and voice commands) and giving him the perception of the elements with which he can interact using feedback (visual and auditory) and feedforward.

Keywords: Mixed Reality, Microsoft Hololens, Vuforia, Unity, Education

1 INTRODUCTION

The Mixed Reality (MR) is a fusion between physical and digital world that makes possible new interactions between human, computer and environment. While in the Virtual Reality (VR) the user is immersed in a synthetic world having the feeling of being teleported to a different reality from that one that he is living, in the Augmented Reality (AR) it is possible to add digital information superimpose them to the real environment, through the camera's metaphor, and hence with the usage of devices, such as smartphone and tablet.

Instead, the mixed reality is not limited by the camera's metaphor and allows the user to stay in his world by adding elements, called holograms, with which he can interact in real-time. They thus acquire the same value as real objects and are made up of light.

The recent evolution of technology has developed in children a greater interest in a form of videogame entertainment rather than in those activities that once allowed intellectual growth. Among these, reading is certainly an undervalued and sidelined activity.

The project aims to make reading more interesting for children, without omitting the technological aspect but keeping the concept of not digitizing it.

A mixed reality application has been developed in which the child, once worn the Hololens, has the opportunity to see the scenes he is reading directly in the surrounding environment, having also the possibility of being able to interact with the various holograms and obliging him to a first reading in order to learn what are the techniques that will carry him to go on in the history. Precisely because of the concept of bringing the child closer to reading and not immersing him in it, it has been decided to make the use of a physical book necessary instead of a digital book. The sentences that the child will read will correspond to voice commands that activate the representation of the various scenes. The use of gestures, such as Airtap, Hold and their combinations, constitute the main interaction techniques.

A children's book usually consists of images. The use of the Vuforia library exploits these images by making available all those tools that make the experience more characteristic.

2 RELATED WORK

In 2018, Tian-Han Gao et Al. [1] developed an application that uses Hololens and Vuforia, whose goal is to develop a virtual museum to make the user experience more immersive. Like these, other applications have been developed in the field of education and cultural growth such as Guangjun Wang et Al. [2] which have developed an application that through mixed reality and augmented reality teaches children notions about animals, what habits and behaviours, using 3D models and virtual books. This application was developed with Unity and Hololens, choosing as AR library Qualcomm to display 3D models. Keane et Al. [3] patented a system that when reading a book, which can be both physical and digital, show a static image or an animation. This animation is composed of a sequence of images of the recognized target also associating holographic visual aids and showing the words of the history that are read aloud by the user.

Our project, compared to the above works, combines the use of Unity and Vuforia for MR with the use of Hololens. The project requires the use of a physical book whose phrases correspond to voice commands that enable the visualization of scenes (with animated 3D models) read by the user. The images of the book, including the cover,

are components that can be used by Vuforia to provide the user with an even more engaging experience.

3 HOLOBOOK

As mentioned above, the aim of the project is to bring the user closer to reading by showing the scenes that he reads directly in the surrounding environment. In particular, there are three phases:

- A first phase where the user learns the tools necessary to use the application in the best way.
- A second phase where the user chooses the book to read.
- A third step where the user starts reading.

Several input systems have been used in these scenes, including:

- Gesture: input events based on human hands. More specifically, we have used the AirTap and a gesture that we have called Drag, but it corresponds to the combination of the AirTap, Hold and movement.
- Vocal commands: input events based on the user's voice. More specifically, several keywords have been defined which correspond to voice commands that once recognized launch an event.
- Gaze-targeting: a form of input that interacts with the world based on where the user is looking. The gaze interacts with objects in the scene through use of a pointer, called "cursor".

3.1 Phase 1

This phase can be considered as a tutorial that explains to the user the mechanisms on how to interact with the elements he will find throughout the story. Since the type of user this application is aimed at are children, the focus was on how to make the child participate in the explanation through animations and interactions designed specifically to keep the child's attention strong.

The elements in this phase are essentially two: a guide character who has the main task of explaining how the application works to the user and a 3D dialogue box in which the sentences that the character is saying are shown. Each string, made up of a set of characters, appears showing each character that composes it in a scanned manner; in this way the explanation does not take place in an auditory way but through a first reading designed especially for children who do not yet have a fluent reading.

The dialogue box is an element with which the user can interact. To make the child more involved, the character will ask him questions with two possible answers, as shown in fig 1.

Figure 1: Guide character with 3D dialogue box where interaction is allowed.

To answer, the user must first gaze-targeting one of the two answers. Once one of them is focused, visual feedback (green color) is returned, and only then it will be possible to perform the AirTap to select the focused response, as shown in fig 2.

Figure 2: Feedback after gaze-targeting and relative feedforward so as to perform the gesture to select the answer.

Since it could be difficult to understand which holograms the user can interact with, it has been necessary to insert feedback to give him a perception that allows him to understand which of the holograms are interactable. In this way best practices are respected [4]. The previous feedback obtained from the focus on the writing is called "hover", which constitutes a state obtained from the transition from the observation state to the focus state.

Together with the visual feedback, the user is helped in the execution of the gesture through the feedforward which shows exactly the movement that the user must make with his hand in order to interact correctly.

From a theoretical point of view, this dialog box is nothing more than a *dialog tree*: the various answers are associated with different animations and strings activated once the AirTap has been carried out on one of these answers. In some cases the animations have also been used to identify the various support elements to the user, as shown in fig 3.

Figure 3: The guide character points to the suggestion box.

The guide character indicates to the user a box located at the top right that will appear when he will need some suggestions to continue in the story. This suggestion box is nothing more than a canvas anchored at the top right which will therefore maintain the same position even when the user moves his gaze. Each time this box appears, an audio feedback is used to announce its appearance. This feedback also improves user perception by providing a better cognitive experience.

3.2 Phase 2

In this second phase, the user will have learned all the tools necessary to use the application. The next step is, therefore, to choose a book. The user, wearing the Hololens, and turning his gaze towards the cover of the book, will see a box trailer appear above it in order to get an idea of the story he will read. To achieve this, it has been necessary to use an image targeting library such as Vuforia, which uses Computer Vision technology to recognize and track images and 3D objects in real-time. Vuforia consists mainly of 3 elements:

- Images Database
- AR Camera
- ImageTarget

Within the Vuforia site [5] it is necessary to create a database containing the images to be recognized. Vuforia will return a license key to be introduced into the application. This license key allows us to download the database in Unity format.

The ImageTarget object is associated with the image to be recognized through the AR Camera, and once recognized, the user will see the associated hologram appear. Fig 4 shows how the hologram (story trailer box) appears above the book cover.

Figure 4: Recognized book and trailer box attached.

The user has the possibility to interact with the trailer box by performing the AirTap on the various buttons (Play, Pause and Resume). These buttons also have visual feedback returned during gaze-targeting. In case the user does not want to see the trailer, he can dissolve it by making a Drag-type gesture, explained to the user through the feedforward, as shown in fig 5.

Figure 5: Gaze-targeting on the Play button and feedforward of the Drag gesture.

The application will recognize if the user has already seen the trailer and will allow him to choose whether to view it a second time or not. For the same reason, he can call up a menu through the voice command "Menù" that allows him to disable (or enable), through the combination of gesture and gaze-targeting, the suggestions of the guide and/or the various feedforwards, such as shown in fig 6. Similarly, it is possible to use the voice command "Esci" to dissolve the menu.

Figure 6: Menu with the first suggestion enabled and the second not.

3.3 Phase 3

In this third phase, the user will start reading the book. The phrases pronounced by the user while reading correspond in part to voice commands which, once recognized, enable events including the appearance of the 3D characters in front of him and the activation of some animations representative of what he has just read. In addition, the Vuforia library has been used to signal the start of a particular animation to the user. The library labels the book images as ImageTarget and shows above the page that the user is reading an animated arrow pointing forward with the text "Animazione in corso!", as shown in fig 7.

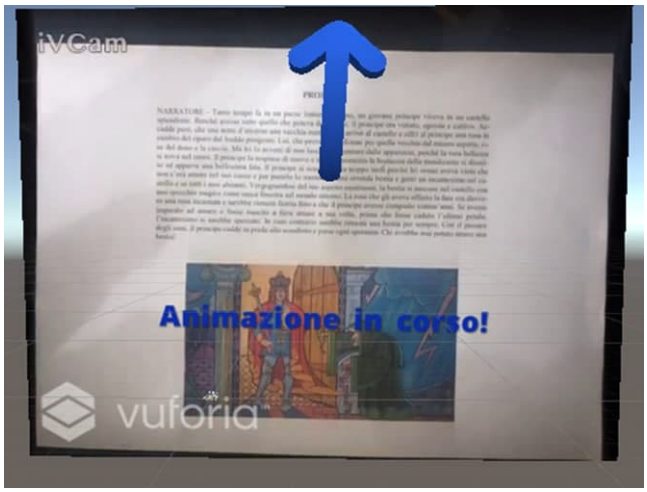

Figure 7: Animation in progress indicated via Vuforia.

The holograms projected above the sheet of paper are anchored on it and so, if the user moves the sheet of paper, the holograms are moved with it. The animations are activated only when the user has finished reading the sentence containing the keywords so as not to interrupt the reading several times and lead the user to have to read the sentence again if necessary to understand the story better. The user will understand that he can interact with the game objects by performing gaze-targeting and checking if the guide character's head appears above them, as shown in fig 8.

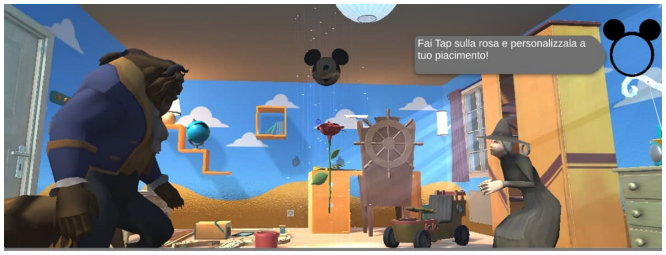

Figure 8: Representation of a scene read by the user in which the guide character's head indicates the possibility of interaction with the rose.

The user could interact with the game objects through the gestures he has learned in the previous phases, and he will always be guided by the suggestion canvas which can be deactivated by recalling the menu through voice commands.

4 GENERAL ASPECTS

Several important aspects have been followed in order to improve the user experience:

- Positioning of holograms: they are positioned within 2m of distance from the user and therefore they respect best practice [6].
- Billboard and Tagalong: these are two behavior concepts that have been applied to some mixed reality objects. In particular:
  - Billboard: the hologram is always facing the user even when he moves to a different position [7]. This behavior has been associated with the guide character, the dialog box and the menu called up via voice command.
  - Tag-Along: it has been used to keep objects in a range that allows the user to interact comfortably (2m). From a practical point of view, the Tag-Along allows the object to follow the user's gaze, always positioning itself within his visual field in a central position [8]. This behavior has been applied to the menu.
- Design: the graphic elements such as the canvas, the dialog box and the trailer have been designed following the design guidelines [9].
- Animations: the various characters in the story have been animated using Mixamo [10]. For one character in particular it has been necessary to use Blender [11] to modify the skeleton and make the animations more fluid. In addition, animations have also been applied to objects, for example, when they have to appear in the scene. In this way, the user has the feeling of adding a new element to his room.

5 CONCLUSIONS AND FUTURE WORKS

Holobook is a Windows Holographic Platform application that was developed on Unity, with the aim of bringing the user closer to reading by showing the scenes he reads in real-time. For this purpose, we focused on the interaction between user and holograms through a guided learning on the different types of interaction, respecting the best practices of Hololens and working for facilitate the perception that the user acquire during the story.

At the moment, the project has been developed using a room as a 3D model to simulate the behavior of the characters in the real environment and develop several important aspects such as the perception of the surrounding environment by the character. To this end, it was ensured that

the characters did not sink and that they turned towards the user if the latter changed position in the room.

Among future developments, a goal is to replace the prefab of the 3D room with a real room mapped using SpatialMapping, and once the meshes of the various surfaces have been obtained, perform SpatialUnderstanding to understand which type of surfaces are and which of these allow positioning of the characters. More specifically, what we want to achieve is that these surfaces do not occlude the character, thus making his positioning "intelligent". For this purpose, a first version has been implemented using a simulator by taking a pre-mapped room through SpatialMapping in which SpatialUnderstanding is carried out so as to understand what could be the most suitable position to place the holograms based on their characteristics. In addition to an intelligent positioning of the characters, SpatialUnderstanding also helps us to use the elements already present in the real room. For example, if among the objects in the real room there was a table or an object on which to place something, the SpatialUnderstanding would detect it as a "Platform" type surface and it would be possible, in the case of our project, to remove the table present in fig 3 and use in its place the table of the real room to place the 3D house always present in fig 3.

In addition, another goal is to use Hololens to test application scenes that implement the Vuforia library because it is not compatible with the emulator.

Another purpose is to give the user a role. Using the dialogue tree, the user could play a different role based on his choices during the story. For example, he could help the characters in the story or becoming an antagonist. All this, however, without distorting the initial objective of the project, which is to bring the user closer to reading physical books.

References

[1] T.-H. Gao, Q.-Y. Tian, and S. Zhou, "The virtual museum based on hololens and vuforia," *4th Annual International Conference on Network and Information Systems for Computers*, 2018.

[2] G. Wang, Y. Qian, Y. Zhang, Z. Lu, X. Chen, and D. Liu, "Design and implementation of children's games based on mixed reality," *IEEE International Conference on Computer Science and Educational Informatization (CSEI)*, Aug 2019.

[3] B. E. Keane, B. J. Sugden, J. Robert L. Crocco, C. E. Miles, K. S. Perez, L. K. Massey, M. J. Lamb, and A. A.-A. Kipman, "Synchronizingvirtual actor's performances to a speakers voice," U.S. Patent 9 035 955 B2, May 2015.

[4] Microsoft, "Interactable object." `https://docs.microsoft.com/it-it/windows/mixed-reality/design/interactable-object`, Jun 2019.

[5] Vuforia. `https://developer.vuforia.com/`.

[6] Microsoft, "Hologram render distances." `https://docs.microsoft.com/en-us/windows/mixed-reality/develop/platform-capabilities-and-apis/hologram-stability#hologram-render-distances`, July 2020.

[7] Microsoft, "What is billboarding?." `https://docs.microsoft.com/it-it/windows/mixed-reality/design/billboarding-and-tag-along#what-is-billboarding`, Mar 2018.

[8] Microsoft, "What is a tag-along?." `https://docs.microsoft.com/it-it/windows/mixed-reality/design/billboarding-and-tag-along#what-is-a-tag-along`, Mar 2018.

[9] Microsoft, "Designing with color." `https://docs.microsoft.com/en-us/windows/mixed-reality/design/color-light-and-materials#designing-with-color`, Mar 2018.

[10] Mixamo. `https://www.mixamo.com/`.

[11] Blender. `https://www.blender.org/`.