Czech University of Life Sciences Prague

Faculty of Economics and Management

Department of Information Engineering

Bachelor Thesis

Analysis and comparison of Software Development Kits for Augmented Reality

Author: Mussin Assyltas

Supervisor: doc. Ing. Vojtěch Merunka, Ph.D.

CZECH UNIVERSITY OF LIFE SCIENCES PRAGUE

Faculty of Economics and Management

BACHELOR THESIS ASSIGNMENT

Assyltas Mussin

Systems Engineering and Informatics
Informatics

Thesis title
Analysis and comparison of Software Development Kits for Augmented Reality

---

Objectives of thesis

The main objective of this thesis is to design and implement Augmented Reality (AR) applications using frameworks and software development kits to conduct a comparison of their performance and usability.

Methodology

The approach to this thesis will be based on a literature review for the theoretical part and the creation of applications in the practical part.

The first part is theoretical in nature and is based on a review of the literature. Numerous sources, such as books, articles, and research papers, will be utilized to gather information on augmented reality and software development kits. This will become a basis for the practical part, allowing the author to identify areas of comparison of individual software development kits.

On the second, practical part, applications will be designed with the help of AR SDKs. The programs will be used to track and identify objects. With these common purpose applications, it will be possible to analyze and compare the performance and usability of various SDKs.

The proposed extent of the thesis
25-40 pages

Keywords
augmented reality, software development kits, tracking, comparison, applications

Recommended information sources

D.Schmalstieg & T. Hollerer. Augmented Reality: Principles and Practice, Addison-Wesley publishing, ISBN-10: 0321883578, ISBN-13:978-0321883575, 2016

Erin Pangilinan; Steve Lukas; Vasanth Mohan. Creating Augmented and Virtual Realities. O'Reilly Media, 2019. ISBN 10: 1492044199, ISBN 13: 9781492044192

Handbook of Augmented Reality, Springer, New York, NY, 2011. ISBN-13: 978-1493901272 ISBN-10: 1493901273

Steve Aukstakalnis. Practical Augmented Reality: A Guide to the Technologies, Applications, and Human Factors for AR and VR (Usability) 1st Edition. Addison-Wesley Professional, 2016. ISBN-10: 0134094239, ISBN-13: 978-0134094236

Expected date of thesis defence
2021/22 SS – FEM

The Bachelor Thesis Supervisor
doc. Ing. Vojtěch Merunka, Ph.D.

Supervising department
Department of Information Engineering

Electronic approval: 16. 3. 2022

doc. Ing. Jiří Vaněk, Ph.D.

Head of department

Electronic approval: 16. 3. 2022

doc. Ing. Tomáš Šubrt, Ph.D.

Dean

Prague on 16. 03. 2022

Declaration

I declare that I have worked on my bachelor thesis titled "Analysis and comparison of Software Development Kits for Augmented Reality" by myself and I have used only the sources mentioned at the end of the thesis. As the author of the bachelor thesis, I declare that the thesis does not break any copyrights.

In Prague on 16.03.2022 _____

`

Acknowledgement

`

Analysis and comparison of Software Development Kits for Augmented Reality

Abstract

The main purpose of this thesis is to conduct a comparative analysis of software development kits of augmented reality. Also, to familiarize with the topic of augmented reality and develop the applications.

The thesis is divided into two sections: theoretical, where the primary objective is to conduct a literature review in order to gain a deeper understanding of augmented reality technology and software development kits; and practical, where the primary objective is to conduct the aforementioned comparison and determine which SDK is best for the average user.

Keywords: Augmented reality, software development kits, tracking, comparison, application.

Analýza a srovnání sad pro vývoj softwaru pro rozšířenou realitu

Abstrakt

Hlavním cílem této práce je provést srovnávací analýzu softwarových vývojových sad rozšířené reality. Také se seznámit s tématem rozšířené reality a vyvíjet aplikace.

Práce je rozdělena do dvou částí: teoretická, kde je primárním cílem provést rešerši literatury za účelem hlubšího porozumění technologiím rozšířené reality a sadám pro vývoj softwaru; a praktické, kde primárním cílem je provést výše uvedené srovnání a určit, která sada SDK je pro průměrného uživatele nejlepší.

Klíčová slova: Rozšířená realita, sady pro vývoj softwaru, sledování, srovnání, aplikace.

Table of content

List of Figures

`

List of Tables

List of abbreviations

AR augmented reality

SDK software development kit

UX user experience

NFT nature feature tracking

SLAM simultanious localisation and mapping

GPS Global positioning system

HMD head mounted display

HHD hand held display

1 Introduction

By their nature, humans are inquisitive creatures. They are captivated by their surroundings, which motivates them to study and discover ways to incorporate parts of nature into their everyday lives. Additionally, they are beings capable of dreaming about something novel and strange. After two million years of progress, we have reached the point where our ancestors' wishes and fantasies are becoming a reality. Modern beings have not stopped dreaming as well. We write stories about people or other worlds to incorporate them into our own lives. It is also possible thanks to scientific advancements and augmented reality.

Augmented reality technologies are the key to a fundamentally new human connection with the digital world, becoming more critical in the global economy and social interactions. AR technologies have received the most significant development in entertainment and marketing areas. The most notable example of a smartphone application having an augmented reality feature is Pokémon Go, which received around 75 million downloads in the first 19 days following its introduction in the digital store, but this is not the end, but merely the beginning of its utilization. Products based on AR technology in industrial production, education, healthcare, and consumer services are the most promising in terms of economic impact.

This thesis studies augmented reality, defining it and where it is used. Additionally, it aims to provide information and assessment of SDK for developing augmented reality applications.

2 Objective and Methodology

2.1 Objective

The main objective of this thesis is to design and implement Augmented Reality applications using frameworks and software development kits and conduct a comparison of their performance.

2.2 Methadology

The methodology of this thesis is based on a literature review for the theoretical part and the creation of applications for the practical part. The first part is dedicated to explaining augmented reality and software development kits using sources such as books, research papers, and online articles. On the second, practical part, applications will be designed with the help of AR SDKs. Applications will be compared based on criteria that will be discussed more in work.

3 Literature review

3.1 History of AR

The broad definition of augmented reality, or AR, is the projection of virtual, computer-generated objects into the physical world using additional devices such as mobile phones, tablets, specialized glasses. Perhaps the most famous and best definition was given in "A Taxonomy of Mixed Reality Visual Displays" (Milgram & Kishino, 1994) as part of the Reality-Virtuality continuum and is "enhancing natural feedback to the operator with simulated cues." In "A Survey of Augmented Reality," R.T. Azuma lists three main characteristics of AR:

- **Combination of real and virtual**. AR is distinguished from other visualization approaches by using visuals created in the virtual environment.
- **Real-time interactions.** The generated visuals appear in real-time, and they can be changed by the people who use them and their environment.
- **Three-dimensional registration**. The tracking not only tracks the generated object but also the user for the changes in perspective and the environment around it for cohesive presentation and occlusion.

The first idea of augmented reality first came with American author Frank L. Baum, famous for his novel "The Wizard of Oz." In 1901, he wrote "Master Key," a story about a boy who got magical electronic glasses, a "character-maker." Glasses allowed the boy to see the characters' initials above people and their personalities. As with all science-fiction technology, fantasy can become a reality with enough time.

In 1968, the technology of AR was achieved by Ivan Sutherland and his student, Bob Sproull. It was a head-mounted display (HMD) that was suspended in the air. Due to its size, mass, and how large it was above the user, it was given the name "Sword of Damocles." It projected a simple image from a computer, a cube, at the center of the display. An object's perspective changes depending on the user's head position. It was simple but groundbreaking (Ivan E. Sutherland, 1968).

Ten years later, in 1978, improving Sutherland's and Sproull's designs, Steve Mann created the first portable AR device called "Eyetap," a prototype of Google Glass. Noticeably, it was

heavier and bulkier than its modern counterpart, and most of the hardware was attached to its frame. EyeTap is not necessarily AR in the sense that we understand it. As Steve Mann wrote, "EyeTap allows under computer control, augments, diminishes, or otherwise alters a user's visual perception of their environment. EyeTap can be displayed to the user in such a way that the user perceives what he/she would otherwise have in the absence of the device".

In 1982, a software engineer, Dan Reitan, led a design team that created the first live augmented reality on television. These were not just simple overlaid graphics but instead real-time images coming from the station's multiple satellite radars. In most weather programs, the usage of AR is still the same.

In 1993, Steve Feiner's Knowledge-based Augmented Reality for Maintenance Assistance (KARMA) system was introduced. The primary purpose of the KARMA was to use the opportunity of projecting graphics and text on the surrounding world to explain how to operate, maintain, or repair the equipment without requiring the user to refer to manuals.

In the 2000s, Jun Rekimoto and Yuji Ayatsuka developed the "CyberCode" visual tag to project computer-generated objects into the real world. This technology offers flexibility when designing AR applications. The main advantage of using visual tags is that they help in the estimation of distance, position, and orientation relative to a tag, further increasing the precision of the displayed objects (Rekimoto & Ayatsuka, 2000).

Before, there was no easy way to create software for AR. This library paved the way for a whole new generation of AR researchers to join the community. The first versions of ARToolKit used the previously mentioned visual tag system to display computer-generated objects. Another noticeable feature of ARToolKit is that it is open source. Mr. Kato decided that it was the best way to test the library.

3.2   Application of AR

Augmented Reality is a technology with great potential. It is integrating increasingly into our lives. The result of that is the constantly growing usage of AR in different fields. From basic amusement games to the military to potentially "game-changing" fireman helmets, there is something for everyone.

People are always looking for new ways to entertain themselves, and augmented reality is no exception. One of the most well-known examples is Pokémon GO, which debuted in 2016. As the name suggests, it is based on a popular series of games, but instead of "catching them all" in fictional worlds, this application adds pocket monsters to your surroundings.

In day-to-day life, AR is starting to be more prominent. IKEA's mobile app provides an AR camera to help decide furniture size and placement. New car models include a built-in camera with an indication for reverse driving.

The use of augmented reality in medicine is often on par with medical imaging, opening up new possibilities for the development of new technologies to improve the performance of surgical teams. The human body is a complicated structure, and direct navigation aids and exhibiting patients' bodies are examples of AR applications. The main advantage is that surgical teams can concentrate on the operative site without being distracted by standard visualization techniques (Joao Cartucho et al., 2020).

As an alternative to using video projectors, surgeons can wear HMDs (Kubben & Sinlae, 2019). As an example, they used the commercially available Microsoft HoloLens, and the software created an MRI (Magnetic Resonance Imaging) or CT (Computer Tomography) scan 3D image, which was projected onto part of the patient's body. Unfortunately, the projections could only be used as reference points. Getting real-time images is problematic for reasons such as:

- Constant exposure to CT X-rays may cause serious health hazards.

- Magnetic resonance exams require the removal of all metallic objects as they may interfere with image quality.

Even if traditional visualizations interfere with AR, they can still be used as a treatment for psychological disorders and phobias (Irene Alice Chicchi Giglioli et al., 2015). According to the study, the patient had katsaridaphobia, a fear of cockroaches. AR helped visualize cockroaches, which could be interacted with by squashing them. This method was effective as, at the end of the session, the patient could approach an actual cockroach.

In other spheres of activity, such as firefighting, augmented reality may help in lifesaving. Not only are citizens' lives in danger in a burning structure with impaired vision, but firefighters' lives are also at stake. Thermal cameras were developed to address this; the disadvantage is that carrying one requires you to use your hands. On the other hand, augmented reality may develop further by merging visual data from a helmet-mounted camera into the firefighter's range of view within the mask. Firefighter aircraft use similar equipment, which includes not just visibility through smoke but also monitoring the perimeters of wildfires and transmitting them instantly to fire stations.

Warfare evolves in lockstep with technological advancements. The military has been interested in using augmented reality technology since the technology's establishment. Long before augmented reality became widely accessible to the general public, the military had already begun to incorporate technology into the visors of fighter jet pilots. This is referred to as a "heads-up display" (HUD). Elliot Flight Automation and Cintel produced the first HUD system in 1961. It indicated the aircraft's attitude and the weapon's target symbology for numerous distinct modes of attack. It obviates the chance of losing focus during operations.

Nowadays, night vision goggles use augmented reality to improve eyesight in low-light conditions. As is well known from history and numerous media outlets, green-washed equipment has drawbacks: extended use may induce headaches, and intense light may cause blindness. The updated version solves these shortcomings and enhances them with a new thermal imaging and outlining technique.

Clearly, the purpose of augmented reality was to improve the user's vision, allowing warriors to stay one step ahead of the adversary. Maintaining track of adversaries, their last known position, and unimpeded eyesight in metropolitan locations (Karlsson, Michael, 2015) is no longer a notion, but a reality.
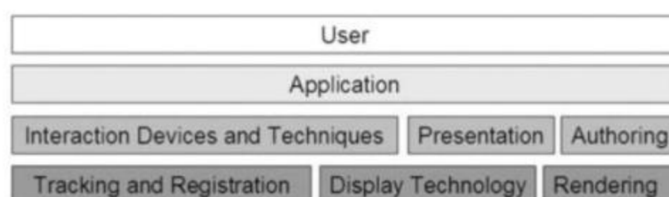

Figure 1. The architecture of AR, Source: O. Bimber, R. Raskar (2005),
Spatial Augmented Reality: Merging Real and Virtual Worlds

16

`

3.3 Technology of AR

The essential components of AR have been established by Oliver Bimber and Ramesh Raskar (2005). A fourth layer is a User, who will evaluate the usefulness of augmented reality. The third layer is dedicated to application development. The second layer is comprised of the tools and devices used to create augmented reality applications. For the last one, the initial layer is constructed using the methods used by augmented reality, which will be explored in further detail in the section following. (Figure 1)

3.3.1 Tracking

Tracking is a complex technique that involves determining the observer's location with respect to the surrounding environment. It is a critical component of the Augmented Reality technology implementation process. The ability of the AR system to distinguish and track the observer from real and virtual things is critical. When an item "fakes" its location in relation to its actual surroundings, it creates an illusion of immersion.
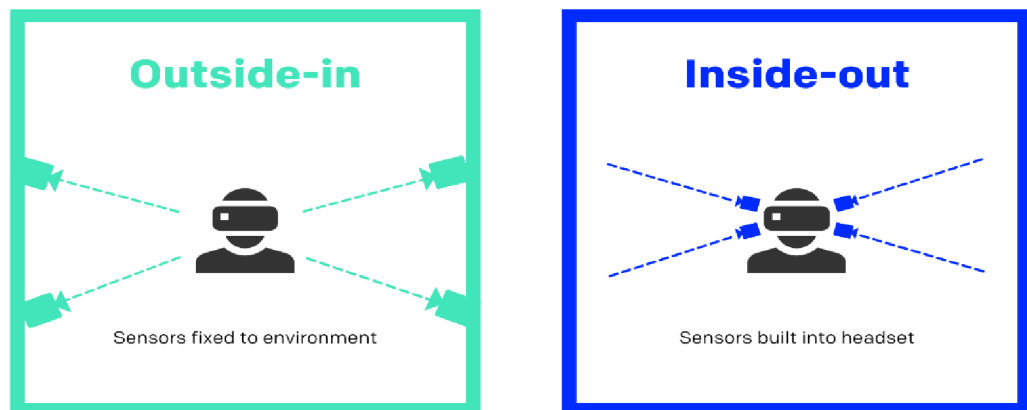
Figure 2. An illustration of Outside-in and Inside-out tracking, Source: https://itigic.com/3d-mouse-how-it-works-and-is-used-in-virtual-reality/

Tracking may be classified into two types: outside-in tracking and inside-out tracking. Outside-in tracking makes use of cameras or other sensors located in a fixed area. In most cases, virtual reality employs outside-in tracking. Inside-out tracking is better suited to augmented reality since it follows the camera's location. The inside-out technique provides more flexibility since it is not constrained by a purposefully built environment, but it takes more computing resources to maintain tracking of the environment.

`

Apart from the two hardware-based tracking approaches, three software-based tracking methods exist: marker and marker-less tracking (Jack C. P. Cheng et al., 2017), as well as location-based tracking.

Marker-based tracking. One of the simplest is marker-based tracking. It requires the downloading of images of the markers as a reference to the application in the production stage. This can almost instantly recognize the marker, and the projected object will be more stable. Unfortunately, this can be a significant disadvantage in this way. It will always be limited by the markers it has in its memory.
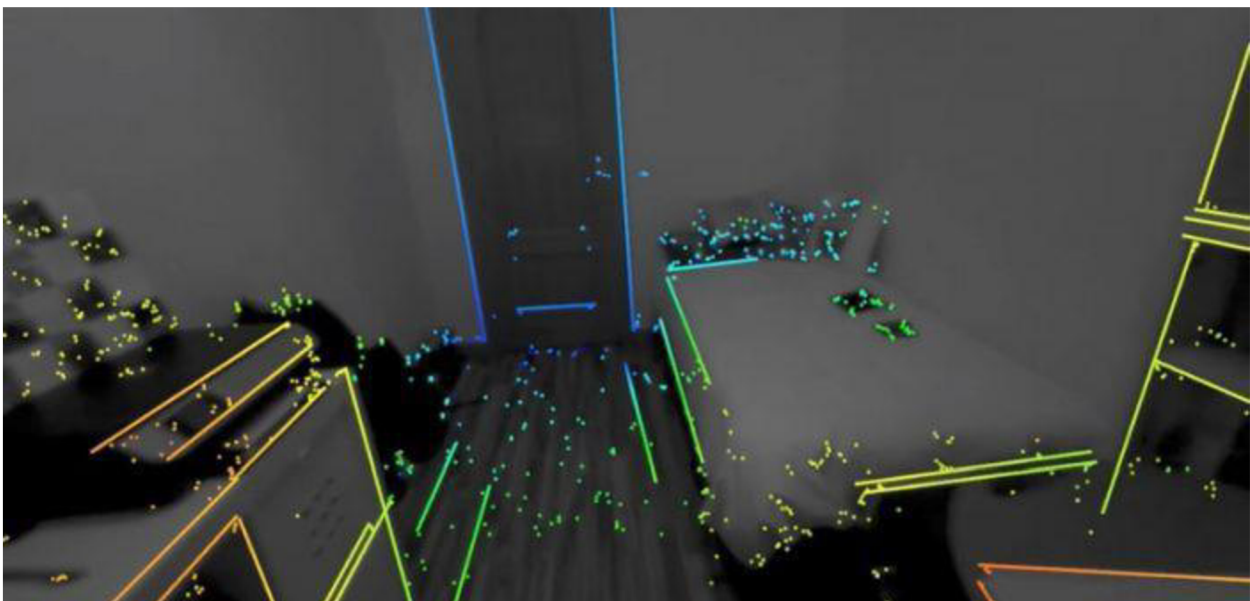
Figure 3. An example of "SLAM" computer vision, Source: https://medium.com/@joodyalmubarak/s-l-a-m-and-tracking-technology-6d1eab3d8a78

Nature feature tracking is a more advanced way of working with markers. Instead of manually placing markers, the system recognizes the main features of the image and creates its own reference points, turning the image itself into a marker. Natural feature-based tracking, on the other hand, is accurate only when the surroundings do not change significantly. When the weather and sun light change, for example, the position of the features moves as well. This makes natural feature-based tracking inherently unreliable in this case, so it doesn't work.

Marker-less tracking. One of the most significant advancements in augmented reality tracking is markerless tracking. AR without markers recognizes objects or distinguishing features in a scene, such as walls or junction points. Identical to the monitoring of natural features, with the exception that it is simultaneous. Simultaneous localization and mapping

occur without any previous knowledge of the surroundings; therefore, the term "simultaneous localization and mapping" (SLAM).

According to Erin Pangilinan et al. (2019), SLAM is a three-dimensional graph or point cloud in which each point represents the position of optical features in the scene. The system can determine the position of a virtual item in the landscape by referencing these points. Thus, the conditional building is composed of walls, windows, edges, and other notable characteristics. A conditional room is a plane with objects contained inside it (floor, ceiling, and walls). Because the algorithm remembers the position of points in space when returning from one room to another, the same points will be in the exact location.

Location-based tracking. Location-based tracking involves the Global Positioning System (GPS) to pinpoint the location of the user and what to be augmented. Usually, this approach also involves additional assistance from other sources such as an accelerometer to understand the position of the device relative to the ground. However, not always will the device have an internet connection available to determine the device's location and assist in tracking. In addition, indoor tracking through GPS might not be accurate enough.

3.3.2 Rendering

The second building block of AR is real-time rendering. It is vital for an augmented reality immersion to understand not just where the object is but also how to portray it. Additionally, it is critical to depict it in such a way that the item seems integrated into the environment rather than just an insert. You must take into account light and shadows, or you may use the rules of physics. A simulated ball, for instance, bounces against a wall. The majority of modern augmented reality systems display virtual items in an enhanced world using usual computer graphics methods. Vulkan API and OpenGL software libraries from Khronos Group are two famous examples.

Despite the study, augmented reality often fails to meet the previously outlined standards for immersion: difficulties with lighting and shadows and managing occlusion (the overlay of the graphical objects on the video feed from the device). The answer was to implement geometry and image-based algorithms, as well as a method known as phantom modeling.

`

3.3.3 Displaying

Displaying is another building block as it represents the central part of AR, augmenting the world around us. That is, altering is not only limited to visual (VAR) but also aural (AAR) and odor-based. For the record, smell-based AR is still conceptual. However, AR is usually associated with visual displays, as they are the most researched. Visual displays are divided into two main parts: monitor-based and "see-through" displays. There is also projection-based spatial AR, which was developed in the following years (Milgram & Kishino, 1994).

Monitor-based AR augments the vision with the use of a monitor (or other sorts of the display). It receives pictures from the camera as input, processes them with the addition of virtual components, and finally displays the result. This form of display has been utilized in television for an extended period of time. The earliest recorded use dates all the way back to 1998 when it was used to denote yardage (measurement of the length or distance, expressed in yards).

See-through displays or HMD and HHD (hand-held display) are like monitor-based displays, and the difference is that they are more affordable and more portable. HMD is a device worn on the head, which gives better immersion as it augments the world in the field of view of the person wearing it. HHD is a device that users can hold in their hands with a screen and enough computational power to overlay additional graphics. There are three forms of HHD: smartphones, tablets, and personal digital assistants (PDA). Smartphones are the most widespread and promising for AR. Smartphones have taken over the entire PDA market, so it is just smartphones and tablets (with pretty vague boundaries between the two). With the widespread use of foldable screens, there may be only phones left. See-through displays can be divided into two groups for different ways to output graphics:

- Video see-through
- Optical see-through.

Video see-through displays are the most common form of augmented reality display since they are utilized in the majority of high-definition televisions because they make it much simpler to change the video feed collected from the camera. Similar technology is used in augmented reality for smartphones.

Optical see-through systems combine computer-generated illustrations with imagery of real world through the curved semi-transparent mirror or the glasses. Microsoft HoloLens and Google Glass serve as a great example.

When the item is shown in the real world, the distinction between optical and video see-through is readily visible. As long as the delay exists, the object's illusion may be disturbed. Due to the fact that calculation and projection are performed by separate components of the device, projecting and viewing the item with a delay behind the HMD might break the immersion. On the other hand, if the delay is present in the HHD, the projection and calculation will be simultaneous since the projection occurs via the device's computations.

Projection-based AR display, or spatial AR compared to other ways of displaying, has more "freedom" as it does not require technology as an intermediary. With Projective displays, pretty much any surface can become an AR display, as projectors project images directly onto real-world surfaces (Zhou, Duh, Billinghurst, 2008). This enables the simultaneous engagement of several users, allowing direct interaction with each other. (Carmigniani & Furht, 2011). However, video projectors and cameras are required to capture and track user movements and display properly. The downside of projection-based AR is that it is not portable and poorly suited for outdoor use as it usually has difficulties with lighting. It overpowers the projected image.

3.4 Software Development Kit

According to with previously described levels of augmented reality building blocks (Bimber & Raskar, 2005), the second layer consists of Interaction Devices and Techniques, Presentation, and Authoring, which the authors avoided expanding the topic further since augmented reality was still in its infancy. Numerous authoring tools in the form of Software Development Kits have evolved during the last 15 years.

Creating AR applications is no simple task for ordinary people. Developers would have to create everything from scratch. It is very time-consuming and resource-consuming. For this issue, software development kits (SDK) have been made. The AR SDK is responsible for the features and components that make up a typical AR application. This includes AR tracking, which represents the visual effects of the program; content rendering,

which relates to how effectively 3D objects can be overlaid in the real world; and scene recognition (Amin and Govilkar, 2015), which makes it possible to recognize, track, and augment feature-rich rooms, scenes, and larger objects.

Four SDKs have been selected based on their popularity: Vuforia, Wikitude, Apple ARKit, and Google ARCore. For the development and working with ARCore and ARKit Unity's AR Foundation has been selected.

Vuforia is used in 50% of AR apps on Google Play and 38% of AR applications on the App Store. Wikitude, on the other hand, is less prevalent, with a 4% App Store rating and a 6% Google Play rating. Over 50% of augmented reality apps use ARKit and ARCore in their own distribution systems for obvious reasons.

Wikitude. Wikitude is a relatively young SDK on the market in comparison to other SDKs. The firm was started in 2008 in Salzburg. Wikitude began as an augmented reality online application. After four years, the business introduced the Wikitude SDK, which quickly became the company's primary offering. It supports tracking of single and multiple images, immediate tracking, 3D model rendering, location-based augmented reality, tracking of objects and scenes, and cloud identification. According to its beginnings, Wikitude enables the production of augmented reality items using its own Wikitude Studio via a web browser. Wikitude is compatible with Android, iOS, and even certain head-up displays. It is offered commercially with a 45-day free trial and an educational license for students and educators.

Wikitude Studio is an authoring tool that enables anybody to build augmented reality experiences simply and without the need for sophisticated technical or programming skills. Additionally, it is used by developers that want to create their own augmented reality experience.

Vuforia. The Vuforia Engine is the most frequently used SDK for augmented reality development, supporting the most popular smartphones, tablets, and eyewear. Developers may simply extend Android and iOS with powerful computer vision technology to build augmented reality experiences that interact realistically with objects and the environment. It is built in C++ and exposes its API to mobile platforms like iOS (C++), Android (Java), and Unity (C#). This is a wonderful option for businesses and brands interested in developing

apps that run on both iOS and Android while avoiding commercial and technical risk. The most recent versions of Unity feature Vuforia as a built-in component. A Vuforia license is necessary to view this information; however, a free developer license is also available.

Apple's ARKit. ARKit is a framework that comes pre-installed with the iOS SDK— developed and presented by Apple in 2017 as a competitor to Google's ARCore. To follow the surroundings, it employs visual-inertial odometry technology, which compares prominent characteristics of the scene frame by frame using data from motion sensors and computer vision analysis from the camera. The result is a high-definition projection. ARKit works with a lot of popular augmented reality engines, like Unity, Unreal Engine, and more.

ARCore. ARCore is the name given to Google's augmented reality SDK. It enables businesses and developers to quickly deploy augmented reality applications on compatible Google smartphones and tablets. Unlike ARKit, it also supports development on iOS devices. ARCore uses a technology called SLAM to establish the phone's position in respect to the surrounding environment, recognizes feature points and planes to comprehend the real-world environment, and provides depth comprehension.

AR Foundation. Unity's augmented reality foundation is a cross-platform development kit built on the ARCore and ARKit technologies from Google and Apple. It enables the development of augmented reality applications for both Android and iOS. Additionally, AR Foundation supports both ARCore and ARKit functionalities. Along with the aforementioned, the AR Foundation focuses on Hololens and MagicLeap development. As a product of a company that is mostly focused on gaming, the AR Foundation is similarly focused on games.

To summarizethe information about AR SDKs the Table 1 has been established.

`

Table 1 Summary of information about SDK features

AR SDKs / Features	Vuforia	Wikitude	AppleARKit	Google ARCore	AR Foundation
Licensing	Free and Commercial	Commercial, with free trial access	Free	Free	Free
Platforms	Android, iOS. Hololens	Android, iOS, Windows	iOS	Android, iOS	Android, iOS, Hololens, MagicLeap
Tracking	NFT, 2D/3D object	NFT, 2D/3D object, geolocation, SLAM	NFT, geolocation, 2D/3D object, SLAM	NFT, SLAM, real-time tracking	NFT, SLAM, 2D/3D object tracking, real-time tracking
Displaying	OpenGl 2.0/3.0	OpenGL 2.0, Vulkan API	OpenGL 2.0/3.0	OpenGL 3.0	OpenGL 3.0

`

4   Practical Part

In the practical part of this thesis, the author will conduct a comparative analysis of several AR SDKs. For comparison purposes, three criteria have been established:

- The first criterion is monitoring many objects with varying geometrical forms and using various tracking techniques, including object and image tracking. The outcomes of this study will demonstrate the adaptability of SDKs: their capacity to adapt to their surroundings.
- The second criterion will emphasize stability. How quickly an item is shown and how fast an object scanned may move before the program stops recording it.
- The third criterion is concerned with the visual presentation or immersion of a shown item. Shadows and occlusion will get special attention.
- The fourth criterion takes user experience into account and addresses the topic of how knowledgeable a user must be to design applications.

The applications will be built in Unity game/graphic engine and run-on Android OS on the Samsung A51 model smartphone. For code C#, a programming language will be used as that's used in Unity. I also want to mention that because of the nature of SDK large portion of work will be done by the SDK itself.

4.1 Applications

4.1.1 Wikitude SDK Application

In order to access the Wikitude SDK first we need complete registration page and get the license key. It is possible to do so in the company's website (https://www.wikitude.com). After choosing the right platform and downloading the plugin for it. The next step is to unzip the file and import it to the new Unity project.
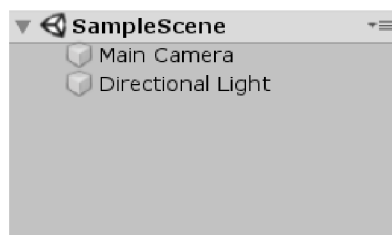
Figure 4. Default starting scene

`

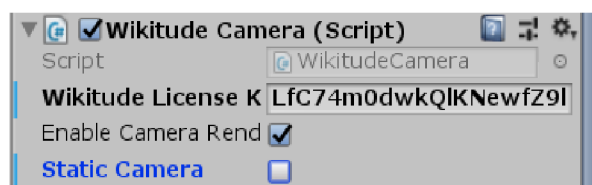*Figure 6. Replacing the Main Camera
with Wikitude Camera*

Figure 5. Inserting the License

After starting the Unity project delete the Main Camera and replace it with WikitudeCamera prefabricated or prefab object and paste the given license key, without it phone camera and SDK will not function. (Figures 6).
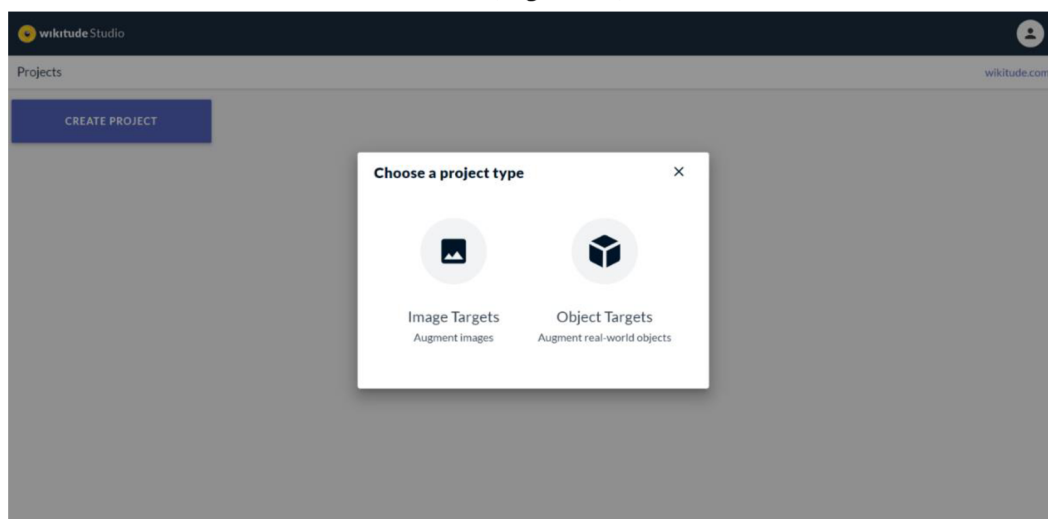

Figure 7. Two options for the targets in Wkitude Studio

The first application is more simple and dedicated to tracking multiple and different dimensional targets (e.g., two-dimensional targets and three-dimensional targets). For this purpose, Wikitude offers its own WebStudio, where previously mentioned features can be integrated into the application (Figure 7). The Wikitude Studio provides options for Image and Object targets. After choosing the first option for image targets and naming the project, it is possible to upload the pictures of the desired object in .jpeg and .png formats. For the two-dimensional marker, a picture of a visiting card from one of the A&O company's hostels will be used as an example and uploaded to the Wikitude Studio (Figure 8). Then it is possible to download markers through email. Open Image Tracker prefab in Unity project

`

Figure 7. Example of an application: building is bound to the card

and upload the marker into it. To display a desired object onto the marker put it on top of the
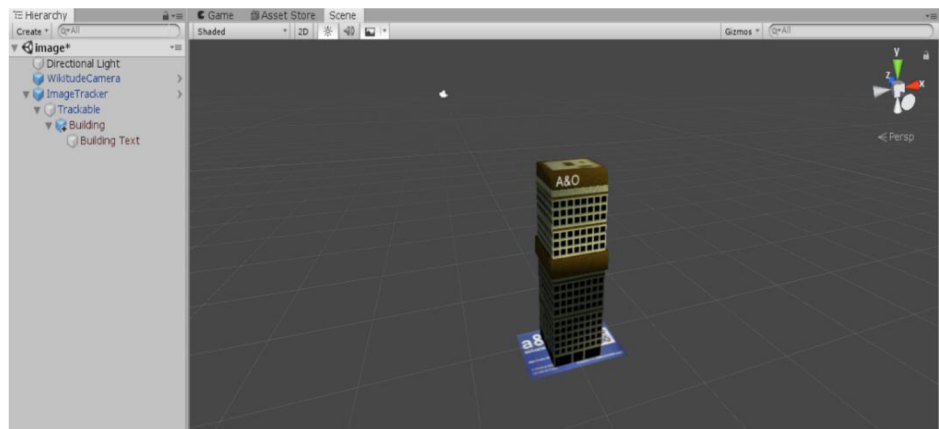image marker

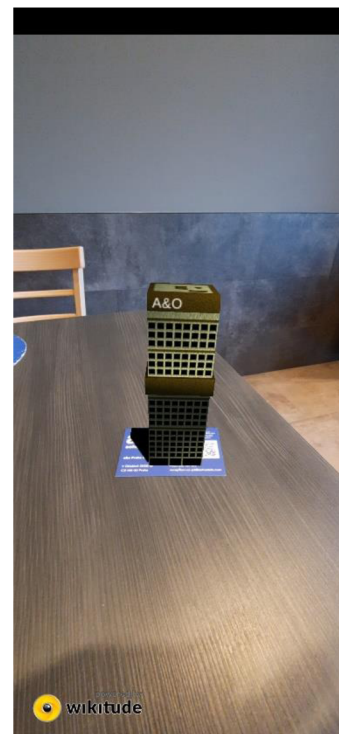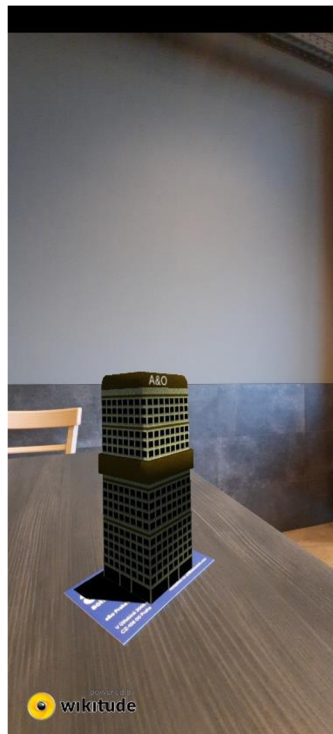*Figure 9. Imported Wikitude Studio target and placement of the object*

Figure 8. An example of the Image target inside the Wikitude Studio

27

`

Next is creating an application based on 3D tracking, and fortunately, this is also included in Wikitude Studio and is under the "Object Targets "option (Figure 7). This function focuses not only on objects but also on the scenery as a marker.
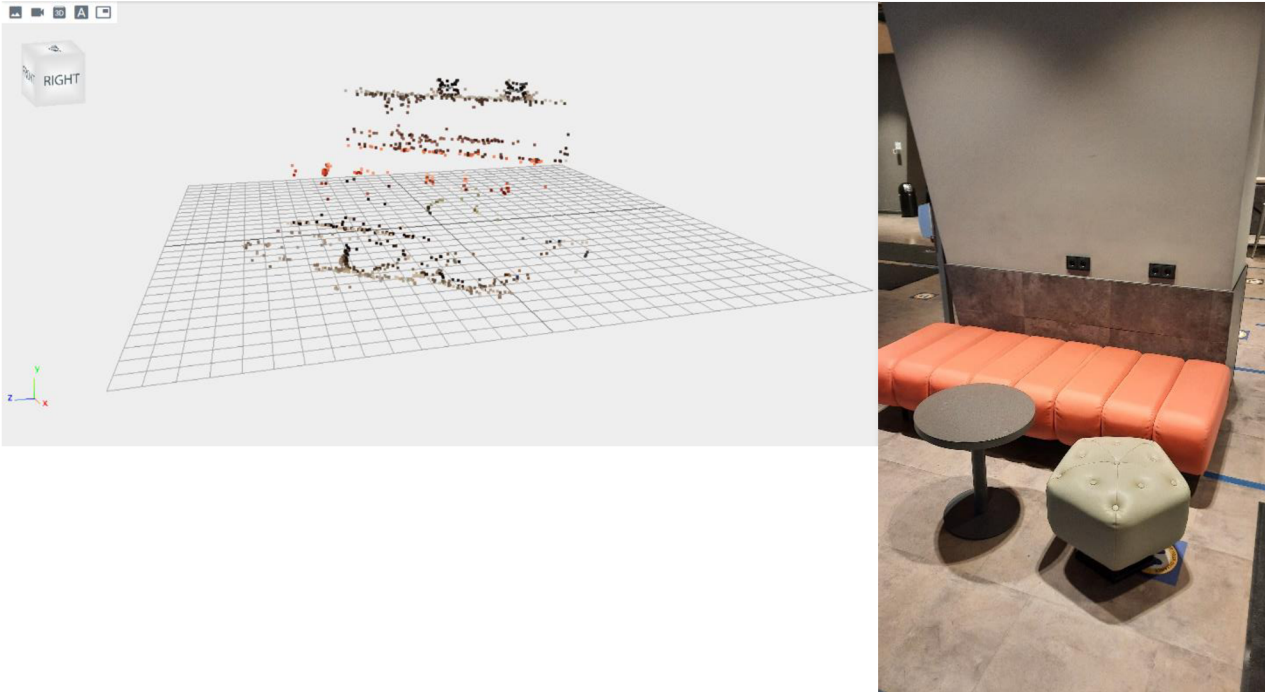

Figure 10. A visualization of point cloud data and the surroundings from which it was derived.

Upload photos of an object or an environment and upload them to the studio. This application covers the environment part of object recognition, as this environment consists of several things itself: two sofas, a table, and a power socket (Figure 11).
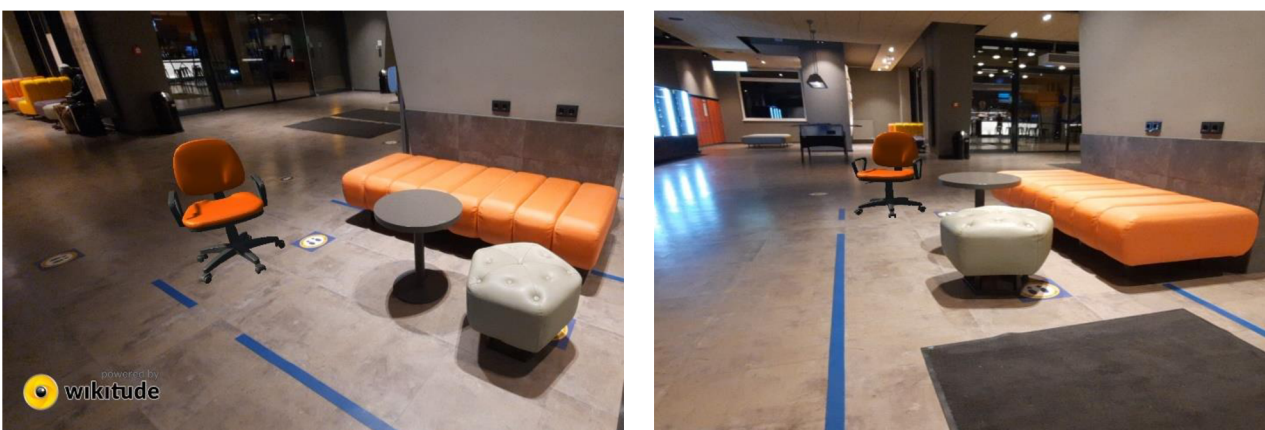

Figure 11. An example of application, the object is placed next to the environment from which the point cloud data has been taken

As seen in Figure 11, the left picture shows the point cloud of the right image; by computing all unique feature points, the SDK visualizes what the camera needs to view to

`

recognize and display the item. Using these points as a reference in space, the displayed object can remain in the same area even from different points of view.

The second designed application has two stages: initialization and Tracking. On the Initialization stage, the application will set the tracking plane on the grid set, which the user has to confirm. Tracking stage, the user can place the object on an already defined plane.

The application has a built-in UI for additional assistance. At the start, the user can set the height of the camera relative to the surface and, by press of the "Initialize" button, fix the grid. Later, two buttons with furniture icons will appear. Pressing the furniture button will show a 3D model of the corresponding furniture piece. A user can adjust their size and position by touching inputs.
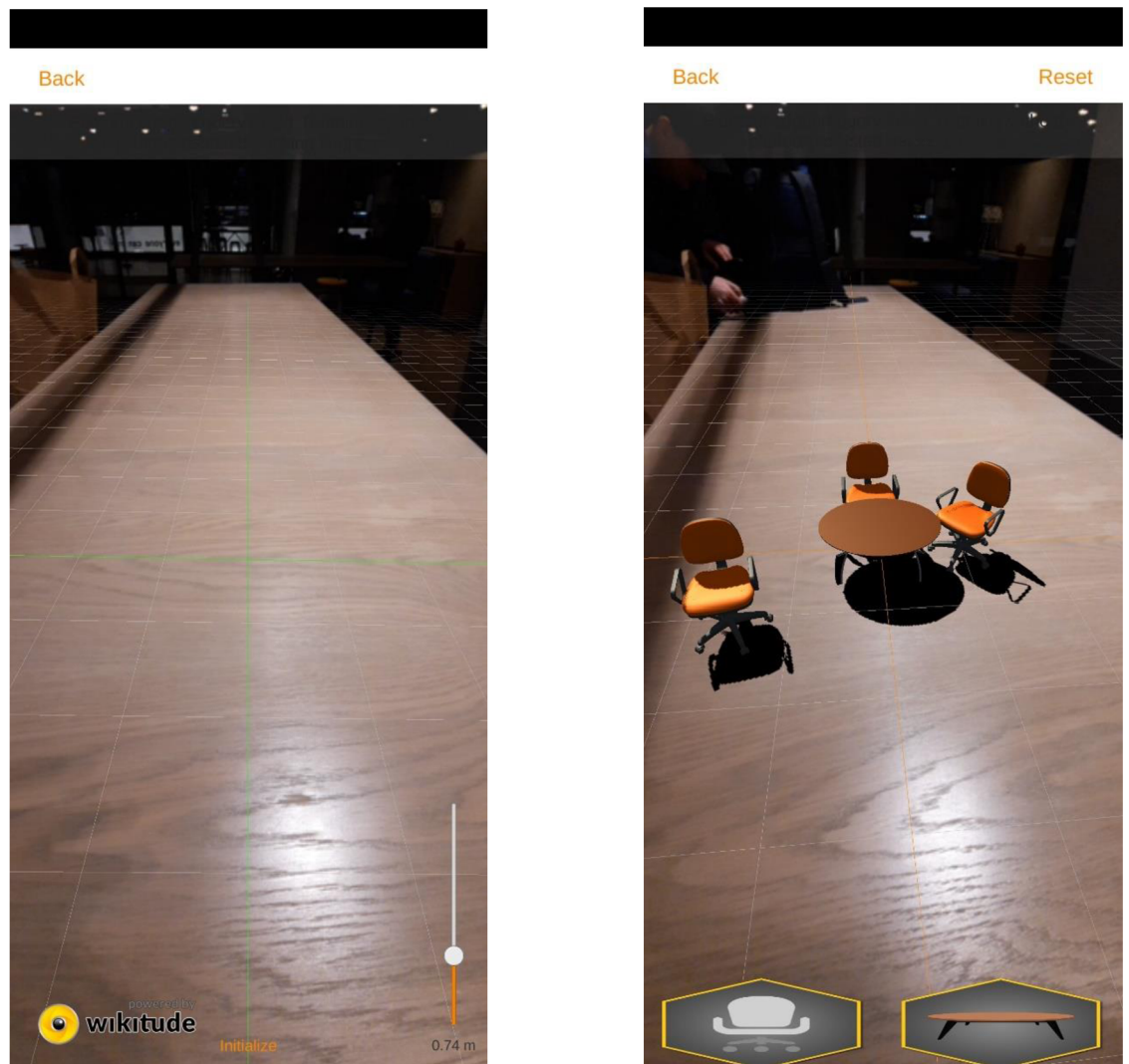

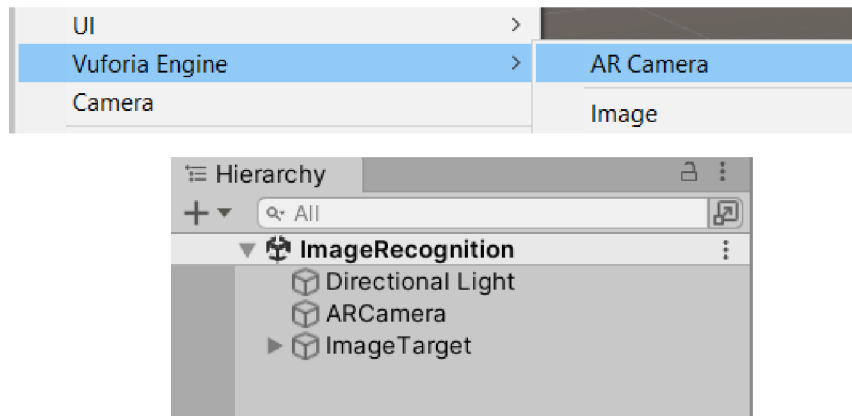*Figure 12. An example of ground detection application*

`

Figure 13. Setting up scene for the Vuforia Engine

4.1.2 Vuforia Engine application

The Vuforia Engine may be downloaded concurrently with the Unity Engine but requires a consent agreement to use it. If it is not included in Unity packages, visit the official website and, after a registration process, the Vuforia Engine can be downloaded and imported into Unity.

The procedure is similar to that of the Wikitude SDK in that the "Main Camera" game object is deleted and replaced with Vuforia Engine's "AR Camera," but a license key is required, and the AR Camera will not function without it.AR Camera, check the Device Tracker option, which will automatically keep track and update the device's location relative to the augmented object and scenery. As with Wikitude applications, Vuforia Engine will assist in designing augmented reality applications based on image recognition and SLAM technologies and attempting to combine them and provide multiple augmentation options. Regrettably, the Vuforia Engine no longer supports object recognition. In recent versions, model and area tracking have been added to the SDK. However, these approaches involve the use of extra complex and specialized instruments.

The first project will focus on Image and Object recognition. Compared to working with the Wikitude SDK, the primary distinction with Vuforia Engine is the degree of integration within the Unity Engine. Rather than visiting a different website and downloading a template for image recognition, this procedure can be completed inside Unity. However, creating image targets inside a website is still accessible.

`

Figure 14. An example of Image target application made using Vuforia Engine

Import the "Image Target" object from the Vuforia Engine assets after the AR camera is in place. The "Image Target Behaviour Script" is available in the Inspector part of the Image Target, and it specifies whether to get the target from the website database or memory. Following the installation of the desired picture as the target and projected object, the Vuforia SDK requests the actual size of the image target. The application is then complete.

For the second application, which focuses on SLAM technology, the Vuforia Engine already has built-in game objects: one for the mid-air manipulations, where objects are locked in 3D space and can be seen as suspended in the air other one is for the ground manipulations. The SDK will calculate and detect the position of the ground and try to place the object onto the detected ground plane. Just ground plane detection will be employed to keep applications equal for the comparison. For this, "Plane Finder" and "Ground Plane Stage" game objects are required. The Plane Finder does the following functions:

- **Anchor Input Listener Behaviour** - Listens to input from the user (such as a tap on the device screen).
- **Plane Finder Behaviour** - Attempts to find an appropriate plane to place content in the real world.
- **Content Positioning Behaviour** - Places content into the real world.

`

The Ground Plane Stage allows your content to be positioned relative to the Ground Plane itself, allowing you to place content relative to the real-world surfaces and scale. Displaying is done by making the object you want to show a child of the Ground Plane Stage. Additionally, the Ground Plane Stage contains an indication to aid with object placement. The indication and the object's size may change the surroundings based on the camera's height and distance. Another advantage of Vuforia Engine's ground plane detection is that the object will stay in its position even if the camera looks away or zooms in.

Figure 15. An example of SLAM ground detection

`

Figure 16. Close up to the one of the placed objects.

4.1.3 AR Foundation application

As with the Vuforia Engine, the AR Foundation is integrated into Unity and is only accessible via the Unity Registry's package management system. It comes pre-installed with the ARCore and ARKit plugins. Additionally, this SDK emphasizes object-oriented programming; it encourages users to customize and choose techniques for specific applications.

As with previous SDKs, the apps developed will emphasize image recognition, object identification, and SLAM. The first step is to construct the scene and import the "AR Session" and "AR Session Origin" objects. The AR Session component is responsible for the augmented reality experience, while the AR Session Origin component is responsible for the augmented reality tracking, as it saves and changes the plane and object's unique points in the Unity scene. One distinguishing feature that separates AR Foundation from other SDKs is that it does not need a license key to access its features.

While the application for the image target is simple, it does need adjustments to the scripts it will employ. First, include the Image Target Manager script in the AR Session Origin, which is already included in the AR Foundation. The next step is to create an image

33

`

reference library that includes the picture used as a target. Then, in the manager's "Tracked Image Prefab" section, add the object that will be shown. As well as add the script which
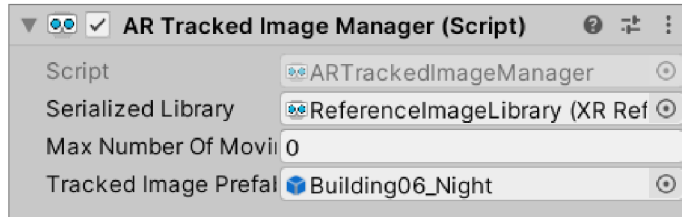

Figure 19. AR Tracked Image Manager, responsible for the storage of image target and displayed object


```csharp
using System.Collections;
using System.Collections.Generic;
using UnityEngine;
using UnityEngine.XR;
using UnityEngine.XR.ARFoundation;

public class ImageRec : MonoBehaviour
{
    private ARTrackedImageManager _arTrackedIamgeManager;

    private void Awake()
    {
        _arTrackedIamgeManager = FindObjectOfType<ARTrackedImageManager>();
    }
    public void OnEnable()
    {
        _arTrackedIamgeManager.trackedImagesChanged += OnImageChanged;
    }
    public void OnDisable()
    {
        _arTrackedIamgeManager.trackedImagesChanged -= OnImageChanged;
    }
    public void OnImageChanged (ARTrackedImagesChangedEventArgs args)
    {
        transform.Rotate(0, 0, 50 * Time.deltaTime); //rotates 50 degrees per second around z axis
    }
}
```

Figure 18. Additional script, responsible for the rotation of the displayed object
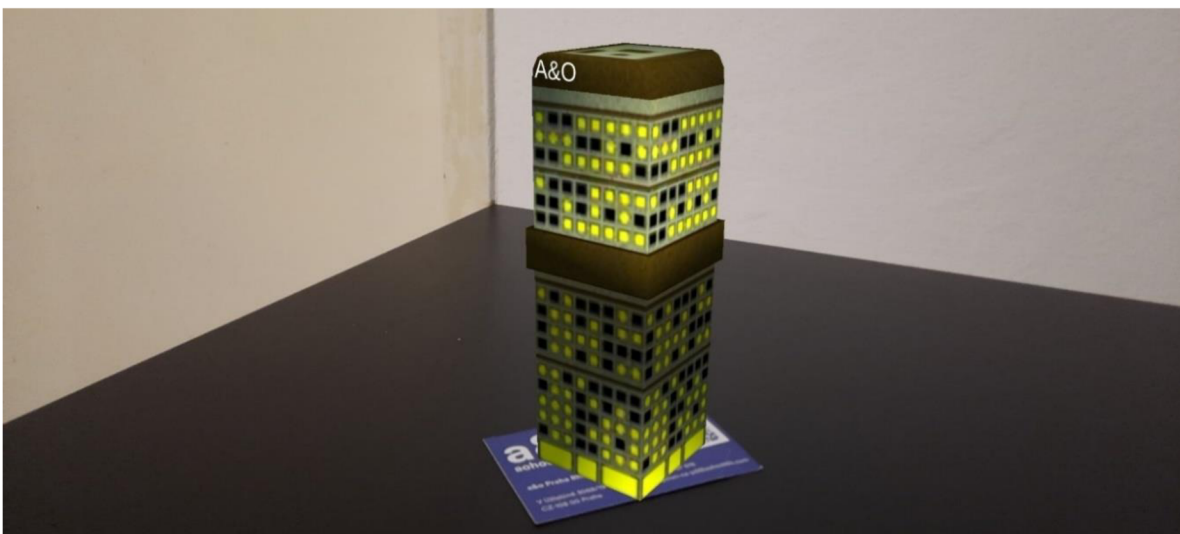

Figure 17. Displayed object rotates around its axis

34

`

will be responsible for the displaying the object.

The detection of objects on the ground is more complicated than in prior applications. While the SDK will aid in development, the user will still need to generate some scripts to operate correctly. First, install the AR Plane Manager script, which will detect and make a new script for the application's functionality. Additionally, create a new game object that will serve as the target.

```csharp
@ Сообщение Unity | Ссылок: 0
void Start()
{
    aRRaycastManager = FindObjectOfType<ARRaycastManager>();
}
// need to update placement indicator, placement pose and spawn
@ Сообщение Unity | Ссылок: 0
void Update()
{
    if (spawnedObject == null && placementPoseIsValid && Input.touchCount > 0 && Input.GetTouch(0).phase == TouchPhase.Began)
        ARPlaceObject();                                        // If the object can be placed, place the object
    UpdatePlacementPose();
    UpdatePlacementIndicator();
}
ссылка: 1
void UpdatePlacementIndicator()
{
    if (spawnedObject == null && placementPoseIsValid)
    {
        placementIndicator.SetActive(true);
        placementIndicator.transform.SetPositionAndRotation(PlacementPose.position, PlacementPose.rotation);
    }
    else
        placementIndicator.SetActive(false);
}
ссылка: 1
void UpdatePlacementPose()
{
    var screenCenter = Camera.current.ViewportToScreenPoint(new Vector3(0.5f, 0.5f));
    var hits = new List<ARRaycastHit>();
    aRRaycastManager.Raycast(screenCenter, hits, TrackableType.Planes);    //By utilizing raycast technology, the target shrinks
    placementPoseIsValid = hits.Count > 0;                                 // or enlarge depending on camera location
    if (placementPoseIsValid)
        PlacementPose = hits[0].pose;
}
ссылка: 1
void ARPlaceObject()
{
    spawnedObject = Instantiate(arObjectToSpawn, PlacementPose.position, PlacementPose.rotation); //Places the object
}
}
```

Figure 20. The script for the ground detection application

Figure 21. An example of Plane detection application developed using AR Foundation

`

4.2 Results

This section will summarize the results from the applications and assess them based on previously defined criteria. Seven applications have been developed for this: three for Wikitude SDK, two for Vuforia Engine, two for Unity's AR Foundation. Each of the applications divulges certain capabilities of representative AR SDK.

4.2.1 First criterion

The first criterion to address is versatility, and the Wikitude SDK is the ideal solution for the average user or developer in this case. While Vuforia Engine and the AR Foundation provide more sophisticated technologies, they also need more professional equipment to work properly. For example, Vuforia Engine's area tracking requires a high-end scanner to scan the surroundings for the information used for the tracking.

4.2.2 Second criterion

Stability is the second criterion. It indicates the object's capacity to be fixed in relation to the environment in which it is positioned. Surprisingly, the Vuforia Engine comes out on top. Despite the familiarity with the AR Foundation's plane recognition project, the Vuforia application showed a tendency for being more fixed on the plane. Even camera motions were ineffective in stopping the augmentation. In the case of the Wikitude SDK, the framework's restrictions cause the objects to vanish or move chaotically if the camera is not stabilized. These are achieved with the extended tracking feature which, upon losing the target, the object will stay on the screen and will track objects similar to the primary target.

4.2.3 Third criterion

The third criterion is one of the main criteria, as the main task of AR is to immerse the virtual world with the real world. The advantage of containing the two most prominent IT companies' SDK did not help the AR Foundation; the main issue was with compatibility of techniques for the immersion is only in some samples of devices. Still, the technology of light estimation makes it superior compared with other SDK.

`

4.2.4   Fourth criterion

The last criterion is also a critical one: user experience. The expectations placed on the user establish the learning curve required to master any tool or application, likewise, with the augmented reality SDKs. With little to no experience working with both the AR SDK and Unity, I determined this would be a critical element for any AR SDK. While working and experiencing the process of developing even simple applications, I discovered that some of my preferences had shifted; while the Wikitude SDK's Web Studio-based development style is simple, it is also limited in what it can do; the AR Foundation development style requires knowledge of the C# language (which is used in the Unity Engine); and despite knowing, the Vuforia Engine is the ideal middle ground between the two. Being integrated into Unity and having a developer portal website gives an excellent service in providing two distinct ways to work. For instance, image targets may be configured in the web studio and the Unity Engine. However, the Vuforia Engine's full capabilities are locked behind a paywall, which brings us to the second problem - payments. Vuforia Engine does provide a free service; however, it is restricted in terms of target count. Wikitude SDK is the company's primary product, only available as a limited-time free trial and has restricted features. On the other side, the AR Foundation is a collection of both free AR SDKs.

`

Table 2. Summary of the criteria, includes the scores.

AR SDKs / Criteria	Vuforia	Wikitude	AR Foundation
Versatility	**Low**, consist of 2D and SLAM., 3D tracking is no longer supported	**Great**, includes 2D/3D and SLAM tracking	**Great**, combination of ARKit and ARCore
Stability	**Great**, the displayed objects are sufficiently steady to approach.	**Low**, since any sudden movement may cause the augmentation to fail.	**Great**, similar to Vuforia, the displayed objects are sufficiently steady to approach.
Immersion	**Low**, presented object is devoid of shadows and occlusion.	**Standard**, although the presented object does not have occlusion, it is compensated for by adjustable shadows.	**Low**, occlusion and light detection options are present only in some devices.
UX	**Great**, offers online studio services, and given its degree of integration with the Unity Engine, enables users to work entirely inside Unity, but is restricted by a paywall.	**Low**, core function of Web Studio work may be its downfall, since it is inconvenient for the integration into the project. Restricted by a paywall	**Standard**, created for Unity, means that all operations can be performed within the Unity Engine but that sufficient knowledge of OOP is required to use its functions.

By summarizing the results on the established table above, it is possible to decide which AR SDK is the best to use. Using the scoring method, it is possible to determine which SDK has the highest score. Three scores have been established: Low, Standard, and Great, which are given 1, 2, and 3, respectively. Overall, Vuforia Engine earns 8 points, Wikitude 7, and the AR Foundation has 9 points, which suggests that the AR Foundation is the most appropriate SDK for its functions.

`

5 Conclusion

This thesis sought to conduct a comparative examination of augmented reality SDKs via application development.

The theoretical section discussed augmented reality technology and its applicability in various sectors, as well as the SDKs. Numerous papers, books, and internet sources have been researched to do this. This part of the thesis covered the history of AR and the technology behind it.

For the practical portion, applications of varying but similar characters have been built with the support of SDKs. The majority of them addressed the fundamentals of each SDK: how to include them in a working build, how to access their capabilities, and the kind of applications that a typical user may create. Four criteria have been defined for comparing AR SDKs: stability, versatility, immersion, and user experience. Occasionally, one SDK outperforms another.

Ultimately, it has been decided that AR Foundation was the best. However, that does not mean that other SDKs are less valuable. Each SDK has its advantages and disadvantages in certain aspects. E.g., Wikitude SDK is built around web development, and it lets users easily integrate the AR to the website; Vuforia Engine focuses on spatial AR and exceeds in it; however, it requires additional tools.

6 References

1. **D.Schmalstieg & T. Hollerer**. Augmented Reality: Principles and Practice, Addison-Wesley publishing, ISBN-10: 0321883578, ISBN-13:978-0321883575, 2016

2. **Erin Pangilinan; Steve Lukas; Vasanth Mohan**. Creating Augmented and Virtual Realities. O'Reilly Media, 2019. ISBN 10: 1492044199, ISBN 13: 9781492044192

3. **Borko Furht.** Handbook of Augmented Reality, Springer, New York, NY, 2011. ISBN-13: 978-1493901272 ISBN-10: 1493901273

4. **Steve Aukstakalnis**. Practical Augmented Reality: A Guide to the Technologies, Applications, and Human Factors for AR and VR (Usability) 1st Edition. Addison-Wesley Professional, 2016. ISBN-10: 0134094239, ISBN-13: 978-0134094236

5. **Alan B. Craig** (2013), Understanding Augmented Reality: Concepts and Applications 1st Edition, ISBN:978-0-240-82408-6

6. **T. Akenine-Möller, E. Haines, N. Hoffman** (2008), Real-Time Rendering, Third Edition, 3rd Edition ISBN: 9781439865293

7. **A.O. Alkhamisi, M.M. Monowar**(2013), Rise of Augmented Reality: Current and Future Application Areas [Online] https://www.scirp.org/html/40277.html

8. **J. Carmigniani, B. Furht** (2011), Augmented Reality: An Overview [Online] https://www.researchgate.net/publication/227164365_Augmented_Reality_An_Overview

9. **P. Milgram, H. Takemura, A. Utsumi, F. Kishino** (1994), Augmented reality: A class of displays on the reality-virtuality continuum [Online] https://www.researchgate.net/publication/228537162_Augmented_reality_A_class_of_displays_on_the_reality-virtuality_continuum

10. **J. Rekimoto, Y. Ayatsuka** (2000), CyberCode: designing augmented reality environments with visual tags [Online] https://dl.acm.org/doi/10.1145/354666.354667

11. **R. Champney, S. J. Lackey, K. Stanney, S. A. Quinn** (2015), Augmented Reality Training of Military Tasks: Reactions from Subject Matter Experts [Online] https://www.researchgate.net/publication/300640705_Augmented_Reality_Training_of_Military_Tasks_Reactions_from_Subject_Matter_Experts

12. **M.C. Juan; C. Botella; M. Alcaniz; R. Banos; C. Carrion; M. Melero; J.A. Lozano** (2005), An augmented reality system for treating psychological disorders: application to phobia to cockroaches [Online] https://ieeexplore.ieee.org/document/1383064

13. https://library.vuforia.com

14. https://developers.google.com/ar/develop/unity-arf/getting-started-ar-foundation

15. https://docs.unity3d.com/Packages/com.unity.xr.arfoundation@4.2/manual/index.html

16. https://www.wikitude.com/documentation/

17. How Augmented Reality is Transforming Retail | Online Digital Marketing Courses. (n.d.). from https://digitalmarketinginstitute.com/en-ie/blog/how-augmented-reality-is-transforming-retail

18. What is Augmented Reality? (n.d.). The Interaction Design Foundation from https://www.interaction-design.org/literature/topics/augmented-reality

19. Early Challenges in AR UX - Library—Google Design. (n.d.), from https://design.google/library/augmented-reality-ux-design/

20. INFOGRAPHIC – AUGMENTED REALITY IN 2016, from https://touchstoneresearch.com/infographic-augmented-reality-in-2016/#:~:text=Augmented%20Reality%20and%20Virtual%20Reality,the%20app%20store%20this%20summer

21. S. Min Ko, W. Suk Chang, Y. Gu Ji (2013), Usability Principles for Augmented Reality Applications in a Smartphone Environment [Online] https://www.researchgate.net/publication/263159230_Usability_Principles_for_Augmented_Reality_Applications_in_a_Smartphone_Environment

22. Optical See Through vs Video See Through on ThirdEye's X1 Smart Glasses from https://medium.com/@thirdeyegen/optical-see-through-vs-video-see-through-on-thirdeyes-x1-smart-glasses-876dda5509d5