Video Calling Through Augmented Reality

Sagar Dhanake[1], Akash Pittalwar[2], Priyanka Dubey[3], Swarneil Pradhan[4]
[1, 2, 3, 4] Dept of Computer Engineering,
[1, 2, 3, 4] Dr. DY Patil Institute Of Engineering and Technology,Ambi, Pune

Abstract- *Augmented reality is an interactive experience of a real-world environment where the objects of real-world are enhanced by computer generated perceptual information, including visual, auditory and olfactory inputs from various sensory modalities. Augmented Reality will play the vital role in enhancing the user experience of better interaction and perception of the output. Augmented reality allows overlapping virtual objects above the images of real objects when seen through your smart device camera. In this article, we have integrated Agora API in a video conference scenario. This article describes the implementation of two scenarios in the video: Integrate AGORA API with live video streaming and Render the live video stream to the AR plane using Agora's Video SDK. We have used AR Core to detect a plane in the room and then make use of Custom Video Source and Renderer function, included in Agora.io Video SDK v2.1.1, to render the live video stream onto the plane. This ends up giving a holographic feel to the video call, just like you see in Sci-fi movies. Agora's video SDK functions RESTful API which is the core of video calling. Agora provides a wrapper for users to easily use WebRTC. Agora has made WebRTC comprehensible to everyone with simplified functions and extensive, instructive documentation. As traditional WebRTC would require you to maintain TURN/STUN servers for relaying data and obtaining IP Addresses respectively, Agora's WebRTC does all of this under the hood, leaving the user with very little to do. The user has no hardware overheads to manage, making the process hassle-free. This drastically cuts down the cost and complexity of implementation. Moreover this application will also concern about multi-user functionality which maybe a revolutionary situation in AR world in near future. So the conclusion of the repertoire of this application is enabling video calling in AR such that the remote people may get exemplary learning and which could be seen as 3D.*

Keywords- Augmented Reality, Agora API, AR Core, RESTful API, WebRTC, Unity Engine, Android SDK, SWIFT UI.

I. INTRODUCTION

Augmented reality (AR) is an interactive experience of a real-world environment where the objects that reside in the real world are enhanced by computer-generated perceptual information, sometimes across multiple sensory modalities, including visual, auditory, haptic, somatosensory and olfactory. AR can be defined as a system that fulfils three basic features: a combination of real and virtual worlds, real-time interaction, and accurate 3D registration of virtual and real objects. The overlaid sensory information can be constructive (i.e. additive to the natural environment), or destructive (i.e. masking of the natural environment). This experience is seamlessly interwoven with the physical world such that it is perceived as an immersive aspect of the real environment. In this way, augmented reality alters one's ongoing perception of a real-world environment, whereas virtual reality completely replaces the user's real-world environment with a simulated one. Augmented reality is related to two largely synonymous terms: mixed reality and computer-mediated reality. The primary value of augmented reality is the manner in which components of the digital world blend into a person's perception of the real world, not as a simple display of data, but through the integration of immersive sensations, which are perceived as natural parts of an environment. The earliest functional AR systems that provided immersive mixed reality experiences for users were invented in the early 1990s, starting with the Virtual Fixtures system developed at the U.S. Air Force's Armstrong Laboratory in 1992. Commercial augmented reality experiences were first introduced in entertainment and gaming businesses. Subsequently, augmented reality applications have spanned commercial industries such as education, communications, medicine, and entertainment. In education, content may be accessed by scanning or viewing an image with a mobile device or by using marker less AR techniques. Augmented reality is used to enhance natural environments or situations and offer perceptually enriched experiences. With the help of advanced AR technologies (e.g. adding computer vision, incorporating AR cameras into smartphone applications and object recognition) the information about the surrounding real world of the user becomes interactive and digitally manipulated. Information about the environment and its objects is overlaid on the real world. This information can be virtual or real, e.g. seeing other real sensed or measured information such as electromagnetic radio waves overlaid in exact alignment with where they actually are in space. Augmented reality also has a lot of potential in the gathering and sharing of tacit knowledge. Augmentation techniques are typically performed in real time and in semantic contexts with environmental elements. Immersive perceptual information is sometimes combined with supplemental information like

scores over a live video feed of a sporting event. This combines the benefits of both augmented reality technology and heads up display technology (HUD).

Figure 1: Example of AR

AR can aid in visualizing building projects. Computergenerated images of a structure can be superimposed onto a real-life local view of a property before the physical building is constructed there; this was demonstrated publicly by Trimble Navigation in 2004. AR can also be employed within an architect's workspace, rendering animated 3D visualizations of their 2D drawings. Architecture sight-seeing can be enhanced with AR applications, allowing users viewing a building's exterior to virtually see through its walls, viewing its interior objects and layout. With continual improvements to GPS accuracy, businesses are able to use augmented reality to visualize georeferenced models of construction sites, underground structures, cables and pipes using mobile devices. Augmented reality is applied to present new projects, to solve on-site construction challenges, and to enhance promotional materials. Examples include the DaqriSmart Helmet, an Android-powered hard hat used to create augmented reality for the industrial worker, including visual instructions, real-time alerts, and 3D mapping. Following the Christchurch earthquake, the University of Canterbury released CityViewAR, which enabled city planners and engineers to visualize buildings that had been destroyed. This not only provided planners with tools to reference the previous cityscape, but it also served as a reminder of the magnitude of the resulting devastation, as entire buildings had been demolished. Amid the rise of data collection and analysis, one of augmented reality's primary goals is to highlight specific features of the physical world, increase understanding of those features, and derive smart and accessible insight that can be applied to real-world applications. Such big data can help inform companies' decision-making and gain insight into consumer spending habits, among others. In educational settings, AR has been used to complement a standard curriculum. Text, graphics, video, and audio may be superimposed into a student's real-time environment. Textbooks, flashcards and other educational reading material may contain embedded "markers" or triggers that, when scanned by an AR device, produced supplementary information to the student rendered in a multimedia format. The 2015 Virtual, Augmented and Mixed Reality: 7th International Conference mentioned Google Glass as an example of augmented reality that can replace the physical classroom. First, AR technologies help learners engage in authentic exploration in the real world, and virtual objects such as texts, videos, and pictures are supplementary elements for learners to conduct investigations of the real-world surroundings. As AR evolves, students can participate interactively and interact with knowledge more authentically. Instead of remaining passive recipients, students can become active learners, able to interact with their learning environment. Computer-generated simulations of historical events allow students to explore and learning details of each significant area of the event site.

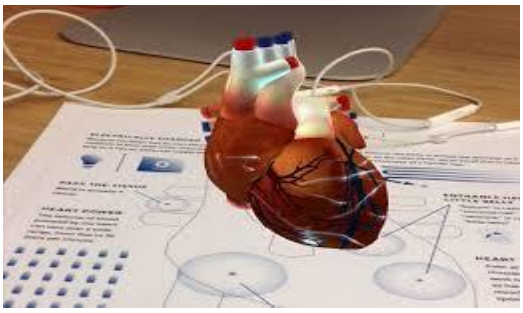

Figure 2 Example of AR

In higher education, Construct3D, a Studierstube system, allows students to learn mechanical engineering concepts, math or geometry. Chemistry AR apps allow students to visualize and interact with the spatial structure of a molecule using a marker object held in the hand. Others have used HP Reveal, a free app, to create AR notecards for studying organic chemistry mechanisms or to create virtual demonstrations of how to use laboratory instrumentation. Anatomy students can visualize different systems of the human body in three dimensions. Using AR as a tool to learn anatomical structures has been shown to increase the learner knowledge and provide intrinsic benefits, such as increased engagement and learner immersion.

II. OVERVIEW

WebRTC (Web Real-Time Communication) is a free, open-source project that provides web browsers and mobile applications with real-time communication (RTC) via simple application programming interfaces (APIs). It allows audio and video communication to work inside web pages by allowing direct peer-to-peer communication, eliminating the need to install plugins or download native apps. Supported by Apple, Google, Microsoft, Mozilla, and Opera, WebRTC is being standardized through the World Wide Web Consortium

(W3C) and the Internet Engineering Task Force(IETF). Major components of WebRTC include several JavaScript APIs: • getUserMedia acquires the audio and video media (e.g., by accessing a device's camera and microphone). • RTCPeerConnection enables audio and video communication between peers. It performs signal processing, codec handling, peer-to-peer communication, security, and bandwidth management. • RTCDataChannel allows bidirectional communication of arbitrary data between peers. It uses the same API as WebSockets and has very low latency. The WebRTC API also includes a statistics function: • getStats allows the web application to retrieve a set of statistics about WebRTC sessions. These statistics data are being described in a separate W3C document. Unity – Unity gives users the ability to create games and experiences in both 2D and 3D, and the engine provides a basic API for writing in C #, for both Unity editor in the form of plugins, and games itself, as well as drag and drop functionality[2]. Before C # became the main programming language used by the engine, it previously supported Boo, which was released with the release of Unity 5, as well as a JavaScript version called Unity Script, released in August 2017, after Unity release 2017.1, in favor of C #. Within 2D games, Unity allows the introduction of sprites with the world's leading 2D provider. In 3D games, Unity allows for texture specification pressures, mipmaps, and resolving settings for each platform supported by a game engine, and provides bump map support, display map, parallax map, screen ambient occlusion (SSAO), shadows powerful using shadow maps, providing texture and full screen processing results. The proposed paper portrays the integration of AR world with video conferencing. In this article, we have integrated an API which explicitly works as video rendering API also known as AGORA API. The Agora Web SDK is a JavaScript library loaded by an HTML web page. The Agora Web SDK library uses APIs in the web browser to establish connections and control the communication and live broadcast services. Agora Software defined real time network, helps us in delivering high quality audio/video streaming performance for the active users over public network. AR Core is a software development kit developed by Google that allows for augmented reality applications to be built. In the perspective of iOS we have SwiftUI developed by Apple that allows to regulate the AR world within iOS. This ends up giving a holographic feel to the video call, just like you see in Sci-fi movies. Agora's video SDK functions RESTful API which is the core of video calling. Agora provides a wrapper for users to easily use WebRTC. Agora has made WebRTC comprehensible to everyone with simplified functions and extensive, instructive documentation This drastically cuts down the cost and complexity of implementation. Moreover this application will also concern about multi-user functionality which maybe a revolutionary situation in AR world in near future. So the conclusion of the repertoire of this application is enabling video calling in AR such that the remote people may get exemplary learning and which could be seen as 3D.Unity and various plugins play vital role in the implementation of the system. AR helps placing the people in the real world who are interacting virtually. Time and monetary constraints may cause ineffective utilization of resources.

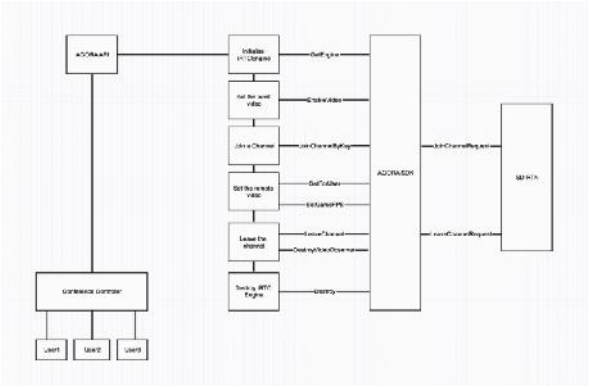

Figure 3 System Architecture

A hybrid tracking and registration technique based on natural features and gyroscopes is used in the proposed system. More efficient remote assistance to people with enhanced feel to immersive interaction. Scope for improvement of robustness and stability of system for wider application. Three dimensional real time tracking registration, Real- time Human-computer interaction and Visual Reality Fusion is combinedly used to explain Combat system in detail. As a result of continuous research and improvement in AR, its application domain has widely spread right from annotation and suggestions to military to medical, film and entertainment. Due to the special nature of AR in various areas and limitation of certain hardware, the cost of development of the system becomes very large. With the help of Mobile AR, 3D models and animation technologies an exergame was developed to tackle worldwide childhood obesity. Smartphone based AR system have solved previous problems of AR system. Unavailability of GPS feature, collision detection between multiple AR objects and less flexible. A device was used to overlay previously captured video onto the real world. Users can freely observe spatial expansion by moving the device. Various inducing techniques were involved to induce user to the same point in the frame from where video was captured. System allows users to experience how camera operators capture scenes by inducing users to move in the same way as camera operators did. Users could freely observe area in the video content as the camera operator did. Scope for appropriate user behaviour where the user action involves switching of scene or translation.

III. LITERATURE SURVEY

Here we present a survey of research papers which has allowed us to identify relevant theories, methods and gaps in the existing research.

- WebRTC role in real time communication and video conferencing. Implemented only on Mozilla where participants' browser did not need the support of any third-party plugins [1].
- Using XR technology and WebRTC API, the system helps participants to communicate in AR world, providing users smooth and beautiful immersive experience [2].
- AR helps placing the people in the real world who are interacting virtually. Time and monetary constraints may cause ineffective utilization of resources [3].
- As a result of continuous research and improvement in AR, its application domain has widely spread rightfrom annotation to military to medical, film and entertainment. Due to the special nature of AR in various areas and limitation of certain hardware. The cost of development of the system becomes very large [5].

IV. ALGORITHMSAND MATHEMATICAL MODEL

Using 3A and an AI-powered noise cancellation algorithm, Agora's platform adapts to variant acoustic conditions toremove ambient and distracting noises, ensuring voices come through crystal clear.

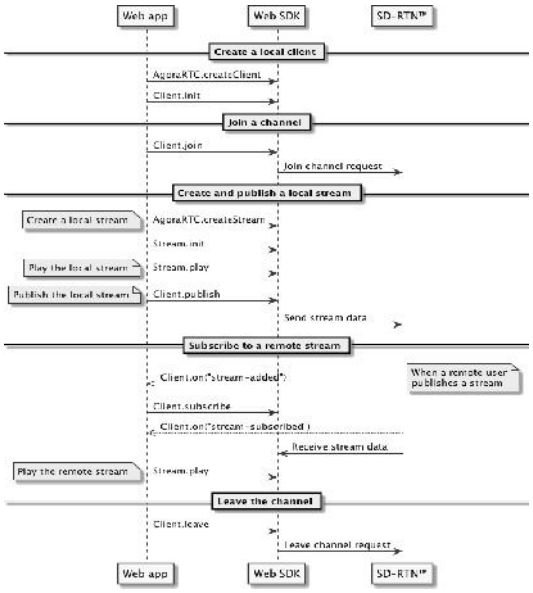

Figure 4 Mathematical Model

V. IMPLEMENTATION AT GLANCE

Figure 5

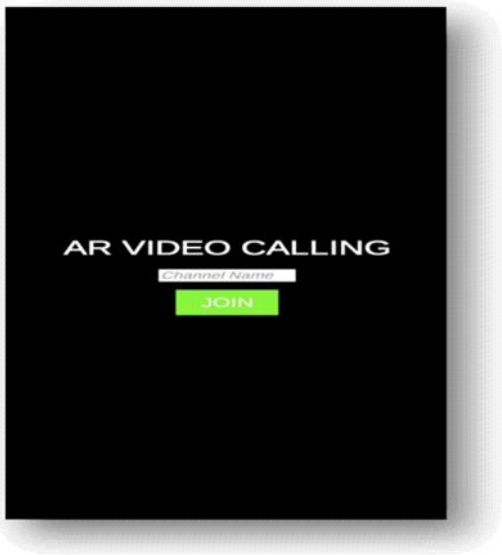

Figure 6

Figure 7

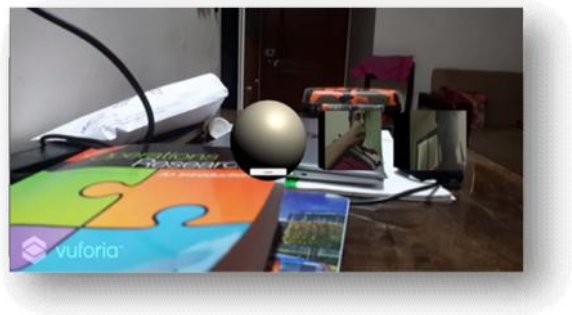

Figure 8

We have thus presented a collaboration of Augmented reality with Video conferencing which focuses on immersive interactions. At first, we have studied various technologies involved in designing video calling system in Augmented Reality. Second, the limitations of previously implemented related systems were studied meticulously and different techniques are proposed to alleviate them. The proposed idea, advantages and applications mentioned above does not fail to present a lackluster thinking and thus will improve much better in future. The limitations/drawbacks which occur during normal video calling like scarce screen size, low bandwidth size are all overcome with the combination of proposed technologies. Augmented Reality may sound daunting at first, but as the developments occur in near future this technology will be our go to thing. The communication issue between remote people or stakeholders as investment people might get mitigated by the introduction of 3D AR video calling. As the communication improves, the results of marketing couldimprove.

Thus stating our application might prove to be the pioneer in upcoming 3D extended technology.

VI. CONCLUSION AND FUTURE WORK

Accordingly, with glancing the functionalities, overviews and enhancements we have concluded that the merging of AR with video calling, the prospects and usage of such technology is incalculable or limitless. With proper measures and adequate comprehensions, this idea or methodology could foresee a formidablefuture.

The future works in this domain of application can further be substantiated as the technology in the near future prevails. One of the many variegated use could be the portrayal of Data Analysis.

The Augmented Reality might prove to be essential in displaying data analysis and execution of final report structures to the clients or the investment committee of particular organization.

The idea of selling a particular analysis by any data scientist could prove noteworthy by using augmented reality 3D video calling apps. Furthermore, this proposed idea could also be merged with the generalization of LMS. By gaining this, the education field will be enhanced with ease in nearfuture.

VII. ACKNOWLEDGEMENT

Our project usually falls short of its expectation unless aided and guided by the right persons at the right time. We avail this opportunity to express our deep sense of gratitude towards all the encouragements and support. At this level of understanding it is difficult to understand the wide spectrum of knowledge without proper guidance andadvice.

Hence, we take this opportunity to express our sincere gratitude to our respected H.O.D. **Prof. Mangesh Manake,** and the Project guide **Prof. Sagar Dhanake** who as guide evolved an interest in us to work and select an entirely new idea for project work. They have been keenly co-operative and helpful to us in sorting out all thedifficulties.

We are also thankful to our subject teachers for their well wishes and inspiration, for their continuous advice and support without which it would have been impossible. Lastly, we would like to express our deep appreciation to our classmates and our indebt ness to our parents for providing moral support and trustingus.

REFERENCES

[1] George Suciu, Stefan Stefanescu, CristianBeceanu, Marian Ceaparu, “WEBRTC role in real time communication and video conferencing”, Kumar Mridul, M. Ramanathan, Kunal Ahirwar, Mansi Sharma, “Multi-user AR application for video communication in virtual space”, IIT Madras 2019.

[2] Dikai Fang, Huahu Xu, Xiaoxian Yang, Minjie Bian, “AR based method for remote collaborative real time assistance from a system perspective”, Springer2019.

[3] WebRTC Official Website.http://www.webrtc.org/.Accessed 25 June 2018

[4] W3C WebRTC 1.0: Real-time communication between browsers.https://www.w3.org/TR/2018/CR-webrtc-20180621/. Accessed 28 June2018

[5] Jun HE, Peng HAN, Huan LIU, Shiying MEN, Lu JU, Pu ZHEN, Ting WANG,” Research and application of AR Technology”, 2017 IEEE, North Information Control Research Academy GroupCo.,Ltd.

[6] Sung Lae Kim, Hae Jung Suk, Jeong Hwa Kang, Jun Mo Jung, “Using Unity3D to facilitate mobile AR Game Development “, 2014 IEEE World Forum on Internet of Things

[7] T. Arakawa, K. Kasada, T. Narumi, T. Tanikawa and M. Hirose, "Augmented reality system for overlaying a scene in a video onto real world and reliving the camera operator's experience," *2013 IEEE Virtual Reality (VR)*, Lake Buena Vista, FL, 2013, pp.139-140.

[8] Johnston AB, Burnett DC (2012) WebRTC: APIs and RTCWEB protocols of the HTML5 real-time web [M]. Digital codexLLC