LUMOZ – A Real Time Augmented Reality Based Tool for Media Production

Wanigasekara.M.P.W.P. A
Faculty of Computing
Sri Lankan Institute ofInformation Technology
Sri Lanka
aroshapubudu@gmail.com

A.Vihanga Nivarthana
Faculty of Computing
Sri Lankan Institute ofInformation Technology
Sri Lanka
vihanganivarthana@gmail.com

R.M.Bawantha Tilan
Faculty of Computing
Sri Lankan Institute ofInformation Technology
Sri Lanka
bawantharathnayaka@gmail.com

Gankanda G.M.J.U
Faculty of Computing
Sri Lankan Institute ofInformation Technology
Sri Lanka
udeshi012@gmail.com

*Mr.*Thusithanjana Thilakarthna
Faculty of Computing
Sri Lankan Institute ofInformation Technology
Sri Lanka
thusithanjana.t@sliit.lk

Dr.Shyam Reyal
Faculty of Computing
Sri Lankan Institute ofInformation Technology
Sri Lanka
shyam.r@sliit.lk

***Abstract*— Globalization has created vast competition among media content productions. Gaining an audience for these products depends on the quality, correctness, timely content and should be produced according to the target audience's comprehension level. The aforementioned factors can be achieved by using new technology in the process of content production. This trend also affects the local news productions. At present, news delivery and other digital media creations are transforming from traditional delivery methods into more advanced technology-based delivery. For this transformation, Augmented Reality and 3D technologies play significant roles which help improving content's attractiveness and correctness. And also, using this new technology can create a new fan base and increase the popularity of the content. The use of these technologies in an industrial base or startup base still provides many challenges. Lack of budget, time consumption, and insufficient technical knowledge are some of the most highlighting challenges. LUMOZ provides a solution for these challenges. This tool provides easier access to point tracking,3D object placement, gesture controls, 3D data visualization, and 3D model library. Simply LUMOZ can be considered as a budget-friendly and time-efficient tool that can handle 3D objects in real-time in a live streaming environment. The user survey conducted to identify the user experiences of LUMOZ has confirmed that this tool is a game-changer to the media industry.**

Keywords: Augmented Reality, LUMOZ, 3D Technology, News delivery, point tracking, object placement, gesture detection, data visualization, Digital Media Production.

I. INTRODUCTION

Years back, content production carried out by traditional delivery methods started with newspapers, articles, and posts. With the technical development of finding signals and wave patterns, radio broadcasting came into the media world. This broadcasting system's focus was to handle radio waves land-based for local delivery and with time satellite-based waves for international delivery. Voice and audio were the modes of delivery in radio broadcasting. Currently, the media is now interwoven with technology. From the modes of black and white visual delivery embedded with audio, Television delivery paved the way for acknowledgment that technology is part of digital media culture [1]. Fast-moving technology has made the digital media world much faster, related, and advanced information delivery. With high computerized graphical advancements and introductions, every digital media content can stream with eye-catching and easily understandable visual elements. Augmented Reality [2], Virtual Reality, and 3D objects have paved the way for creating visual advancements of digital delivery. These technologies are used by many professions such as television producers, News producers, Advertisement creators, Live news reporters, Mobile journalists, social media content creators, and many more now in the world. But not all can assess technology equally.

Mobile journalists and other live media content creators face many challenges [3] for creating content, gathering content information, and delivering mediums. Sometimes, many technical challenges can arise especially when they are traveling while producing the content on a live broadcasting medium. At that moment, carrying heavy machinery or high-performance laptops is not possible. Usually, mobile journalists and reporters carry a regular laptop and a camera. In this situation, digital content creators, might not have the ability or access to good and high-performance graphical tools or software. The current solution of mobile journalists is that the processing of the content and the delivery of the content will be carried out as post-production with more days of editing the content rather than a live production. Simply, the production time of graphical content can consume much time and energy.

Consider a startup entrepreneur with less budget and a tight schedule. The lack of power of affording high-budget equipment or tools can be a result of producing basic and unattractive digital content. Having good quality and graphical content is crucial to get on with the competition of producing content that can attract much audience when they are streaming. Because in the end, the number of viewers or subscribers is the main target of a digital media company or the creators.

Already there are many tools or software and graphic creator companies available in the market for high graphical content creations. Most of the graphical tools in the market, need high capacity or high-performance devices. These devices are expensive and lack mobility. Such tools are currently used by local companies and creators for their news and content productions. The problem among available resources is that they are either expensive or only facilitate with time-consuming pre-production methods.

Pre-production videos can consume much time. With that, timely content cannot reach the audience at the proper time. Therefore, News productions and other live events use the live streaming concept in real-time [4]. In live streaming videos, it is hard to add 3D objects and use technologies like Augmented Reality.

This research focuses on a working solution for the digital and mass media live content creators and live streaming to use 3D graphic content in their broadcasts. LUMOZ provides a solution to the previously mentioned problem by enabling easy, low-budget access for live media content providers to carry out their work with modern technologies like AR and 3D visual objects. LUMOZ design tool includes main functionalities such as Point tracking, Object placement, Gesture recognition, and Data visualization.

II. Backgroung and Literature Survey

Digital media platforms and digital media marketing has a significant rise comparing to other channels of marketing. The future of marketing will depend on how the digital media technologies grow and how far the digital media platforms use modern technologies to provide advanced content. The approach and the use of Augmented reality for digital media is relatively a new concept. Augmented Reality can be considered as the future of digital media marketing. LUMOZ tool provide functionalities like point tracking, 3D object placement, 3D data visualization and gesture detection using the technologies of Augmented Reality and 3D object manipulation. Following are literature surveys done separately for each functionality.

A. Point Tracking

Feature-based pose estimation in video streams [5]. The basic idea of this article is a vision-based marker-less tracking system that aims at the use of real scene features for estimating the pose of a camera. Where they performed using realistic gestures. They have developed a Lucas- Kanade algorithm in the form of a pyramid and integrated it with the Shi-Tomasi algorithm, using an external tracking tactic. They adjust the posture provided by the external tracker to the 3D points, and then create a display engine that covers the line segments that represent the virtual objects in the wireframe to display the results. Identifying and tracking features using computer vision [6] is a must-have in every augmented reality program. Features here are any object contained in the natural environment. Augmented Reality technology has been introduced using a depth map of an image [7]. What is done here is to analyse a video from frame to frame and convert the images into a grey scale depth map and display it in 3D. This is a technology that is also used in Facebook. Here we can see objects as 2.5d, not 3d.

Point tracking is the process of identifying the specific points which are suitable for placing the objects. These points tracked are useful for both still camera and the camera with the movement. Points tracked on a moving camera needs to stay at the same points identified without changing it position. This helps the 3D object which is placed afterwards to hold its position.

B. Object Placement

Based-on feature tracking, inserting virtual objects into a real video stream and estimating the camera from a collection of single camera video frames [8]. The transformation from 3D objects to object projections needs to be identified to insert or alter 3D formats to target video frames. It has been shown that 3D images can be recreated using multiple images from one single camera under defined internal camera parameters without a camera calibration process. The method proposed is based on the simplification of the internal parameter camera model and the use of projection geometry. For enhanced virtual applications, this method is particularly useful for inserting or changing formats into a virtual video stream. It mentions the special object detection technology in Open CV to identify the object here.[9] It describes the special object detection technology in OpenCV to identify the object. Special algorithms and special libraries in OpenCV are being used. Objects can be identified from image pyramids algorithms, geometric descriptor algorithms
Resource reservation is required in a video-on-demand server to ensure continuous delivery. As a result, each storage device (or striping group viewed as a single logical device) can only accommodate a certain number of client access streams.[10] The number of video files that each storage device can store is also reduced. Multiple storage devices can be used in a video server environment for reasons of availability, gradual development, and heterogeneity. As a result, one or more copies of a video can be stored on various storage devices. There could be a load imbalance among the devices because the access rates to different videos are not standardized. In this paper, we propose a dynamic placement policy (dubbed the Bandwidth to Space Ratio (BSR) Policy) that creates and/or deletes video replicas when mixing hot and cold videos to make the most of a storage device's bandwidth and space. A simulation analysis is used to test the proposed strategy. The evolution of video-sharing sites has drawn substantial investments in contextual advertising over the last decade. The information provided by users is used by popular contextual advertisement platforms to incorporate 2D visual advertisements into videos. Current platforms face several technological challenges, including ad integration with occluding artifacts and 3D ad positioning.[11]

C. Data Visualization

Misleading data spread through incorrect graphs are common issues among Data visualization [12]. Data communication through data visualization is very effective tool and easier way to convey the data to the viewers. One of the methods used in the Data Visualization to read data are from the database is using data manipulation through SQL statements [13]. After data is taken from the database, then different tools for data visualization is used. Then the data exploration is another important step in data visualization. But these tools and software's doesn't meet the ends of creating a visual element that can also be shown as a 3D element on the daily programs. Many forms of data visualizations have been used all around the world and in many fields. Data Visualization has improved the quality of displaying statistics and improve the human thinking process [14]. Not only in broadcasting medias, but this has also become of the decision support system to many fields of work. In LUMOZ data visualization, the focus is to create

visual elements such as graphs, charts, and maps easily and converting them into 3D viewed objects which can be used appropriately in live streaming by the object placement function. System already updated with the basic visual element models which can be further modified with time efficiently. User only has to select the graph, enter the values and select the color theme. The 3D object of the visual element will be created and send off to the Object placement library.

D. Gesture Detection

HCI (Human Computer Interaction) has become an inseparable [15] when working with technology. Gesture detection has become a great promotor of HCI. Gesture recognition and manipulations done according to the gestures has made significant advancements in many fields. In the media field, gesture recognition is important in both physical interactions as well as in HCI. When dealing with 3Denvironmant where 3D objects are manipulated, gestures such as hand movement [17] and location are very much important. Hand gestures has made a natural communication between man and 3D world [16]. Hand gestures can simply replace mouse and keyboard movements like zoom in, zoom out, rotate, and position changers. In this article [16], they have proposed a system where hand gestures are recognized in real time based on computer vision. Identifying the presenter's hand gestures and the presenter's location to place and manage the object accordingly. A gesture of the presenter is detected from the original images from the input devices to place the object. A live review of the object will be provided for the presenter as they move their hands or changing their positions. The hand tracking algorithm can be used to detect the presenter's hand gestures. After the tracking, the system labels the gestures and connect 3D object coordinates.

III. Research Objectives

Main objective of this research is to provide a platform for the live content creators to easily create high graphical effects in their live production. With the current situation of the world due to Covid 19, almost all the possible events and medias has become online. Live and preproduction events have become online. With this situation, many content creators have chosen platforms on digital medias such as YouTube and other social media platforms. For them to create interactive and high graphical content for a live streaming we introduce our product LUMOZ. This tool's main objective is to provide real time manipulation of 3D objects in a live streaming environment and help the content creator to take live broadcasting to the next level.

LUMOZ supports for the following sub-objectives.

- Identifying the positions and points where the 3D objects should be placed.
- Adjusting the 3D object position and size according to the camera movement.
- Providing manual 3D object manipulations with functions like zooming and rotating.
- Controlling the 3D object with pre-defined hand gestures.
- Providing a 3D object library where the user can either add and use their own 3D objects or buy available 3D objects from the library.
- Create 3D data visualization objects and use them in real time.

All these features are provided for digital content creators, who are in live streaming fields like News productions, live on social medias, live events, online education, and other quick productions.

IV. Methodology

LUMOZ is a combined system of all the necessary tools which are used for an AR based live production. Unlike other tools which are already in the market [23] [24], our product can be considered as a flexible tool for media productions which consist of low budget and mainly low processing power. This tool is designed for easier use for the live 3D based AR system which can embed into a live production as in its broadcasting phase.

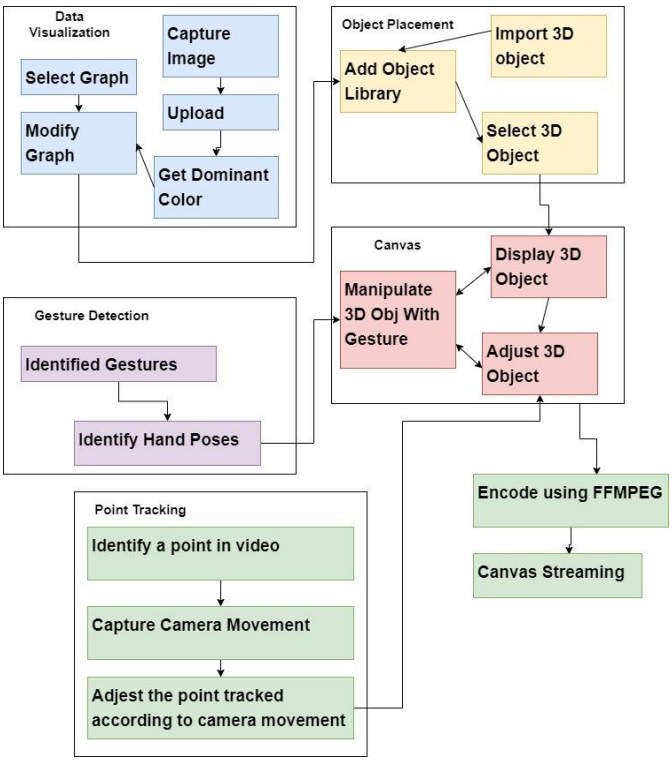

Figure 1:Overall System Diagram

LUMOZ tool basically uses AR technology and 3D object Manipulation. It provides four main functionalities. Point tracking, 3D object placement, gesture detection and 3D data visualization. Figure 1 explains how the four main functionalities blend in to create LUMOZ. All four of these functionalities combine into one canvas which is encoded using FFmpeg and out as live streaming in media platform.

A. Point Tracking

In the point tracking functionality, the point should be identified and adjust the point according to the camera movement. For example, if the point is on the table, even if the camera is moved left to right or vice versa or changes its

angle, the point should remain on the identified location. For this functionality, we use IOT devices to adjust the 3D objects parallel to the camera movement. IOT device consist of an accelerometer sensor which is used to measure the angle and sonar sensor to measure the distance. The concept of this IOT device is shown in the diagram given bellow.

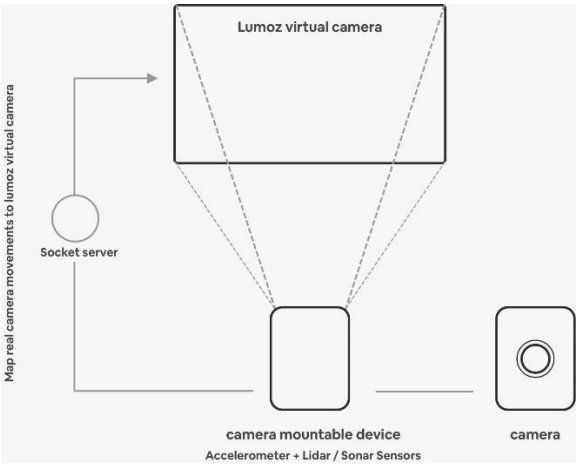

Figure 2: IOT Device

The calculated value obtained using these sensors are then sends to the server to identify the camera location. Then the identified point location is sent to the 3D Object placement functionality.

The process from the canvas to live streaming is shown in the below figure.

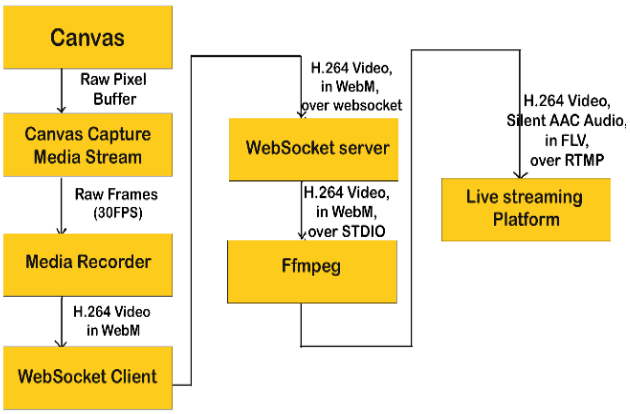

Figure 3: Live Streaming Flow from Canvas to Platform

In live streaming, using web sockets the live API is connected with the Webgl [18] canvas is captured. Then the captured canvas is shown in the live stream. After that it is encoded using FFmpeg and given to the streaming platforms like Facebook, YouTube, etc.

B. Object Placement

3D object placement functionality takes the identified point and uses it to place the 3D object. The main feature of this functionality is the 3D object model library. The user can import 3D models to the library and store them. Node JS is used for importing 3D objects and save them in the local folder. This feature supports file formats like gltf and glb. We do not allow the users to create 3D objects unless they are 3D visual elements. These imported 3D objects then can be used in a live streaming session. We use three.JS for object placement. This functionally helps to keep the 3D object intact to the point and adjust its location and size when the camera is moving. When the camera zooms in, the 3D object should increase size and vice versa. The canvas used in this system contains layers.

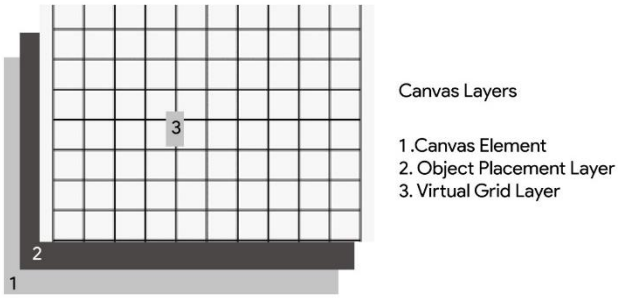

Figure 4: Canvas Layers

The main three layers of the canvas are canvas element which is an HTML canvas, the second layer is the object placement layer which is used to place the 3D object and the third layer is the virtual grid which has the invisible X, Y coordinates. The arrangement of these three layers is shown in the above figure.

C. Gesture Detection

In gesture recognition, the main functionality is to communicate with a 3D object in a live environment. Without any help from another person or a system, the presenter itself has the opportunity of controlling 3D object movement. Gesture recognition mainly focuses on hand movement recognition. This tool already identifies 20 hand landmarks using Media Pipe technology [20].

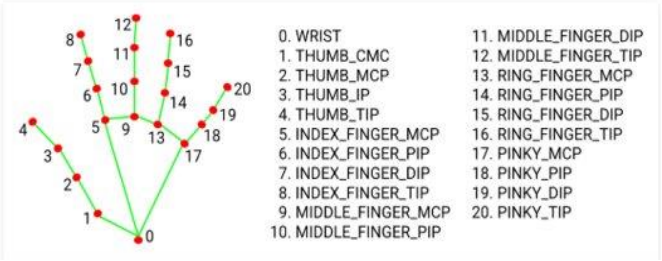

Figure 5: Hand Landmarks

These landmarks [21] are then used for the performance of zooming functions and rotate functions. For the rotate function, the distance between two landmarks is calculated. Using the law of cosine, the angle between the thump and the index figure is calculated for the zooming process. If the angle is more than 0.7 then the zoom function starts. The zoom level is decided by the difference between 2x coordinates for the landmark shown in the diagram below.

Figure 6: Hand Command Rotate & Zoom

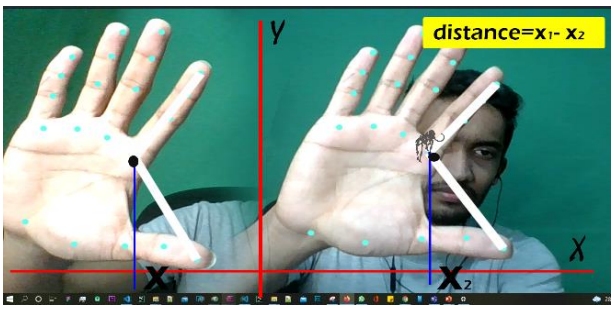

Figure 7: Zoom In/Out Action

D. Data Visualization

The 3D model library used in the object placement component can have 3D objects which are imported using external sources. But for data and statistic representation, charts or graphs should be created in real-time. For that, the data visualization function is embedded into this project to create 3D graphs which can be viewed in a live stream. This functionality supports the user to select the type of graph needed. User can either manually enter the data to the graph or can add a data sheet and generate the graph accordingly. Graph selection and graph color selection are the two main components of this section. The color selection [22] is automated to select the color which is suitable for the presenter's location. For that, an image can be given to the system and the system will identify the most dominant colors of that image and suggest matching colors for the graph [25].

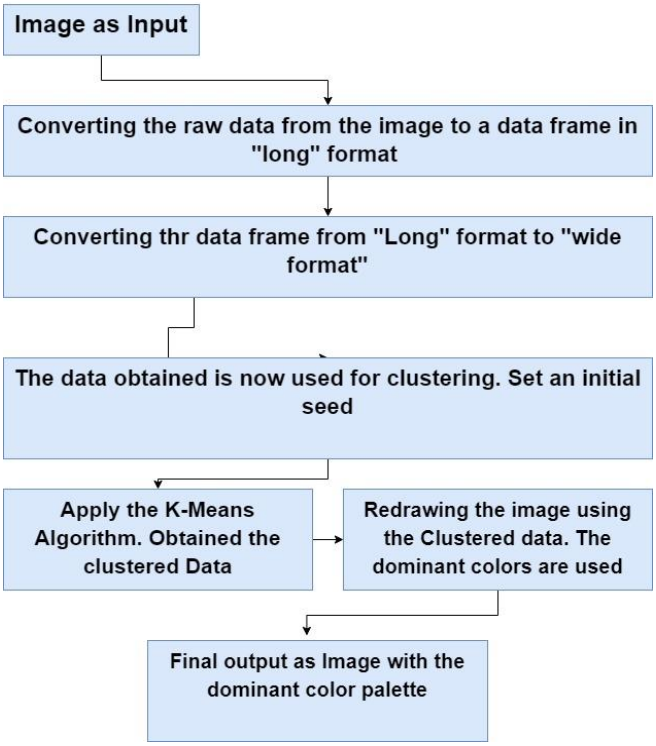

Figure 8: Dominant Color Suggestion Algorithm

V. Results and Discussion

Before starting the research implementation, the information gathering process was carried out thoroughly. By contacting the current professionals of the media field experts, we were able to gather much information regarding the problems of using new technology in content creations. The problems we found were lack of technological knowledge of media field employees, budget limitations, and time management. And the existing tools and software need high-capacity computers and other hardware, so it creates a problem for the startup individuals to afford high graphic supporting hardware. As a result of these findings, we came up with LUMOZ.

Moving forward with the research of creating LUMOZ, we conducted several interviews with the industry experts and individual content creators to get their experiences and knowledge on the topics of live streaming and the use of Augmented Reality. We contacted Ravihara Perera (Assistant Manager Graphics & VFX at TV Derana, Former Senior Executive 3D Generalist & VFX Artist at TV Derana) and Thilini Perera (Senior News anchor & News Producer at Swarnavahini). According to Mr. Ravihara Perera, most of the high-end graphic work is outsourced to external companies. This can cost a huge budget for the company, and it is not possible to use this method for their daily content. And with the interview conducted with Ms. Thilini Perera, we found that they must conduct several rehearsals before using Augmented Reality technology. And during the presentation, the 3D objects are controlled by the controlling room, and they must maintain constant communication between the controlling person and perform according to their commands. She mentioned that it is a very difficult task and needs a lot of time.

As a result of these interviews, we can say that LUMOZ provides a better solution for the use of new technologies like augmented reality in daily content creation without facing any obstacles mentioned by the experts.

LUMOZ provides a budget-friendly and time-efficient live streaming experience with augmented reality for content creators. LUMOZ doesn't need a second person to control the 3D objects. The presenter itself can control the 3D object while presenting. According to Ms. Thilini Perera, self-controlling 3D objects are much easier and effective when conducting a presentation smoothly.

LUMOZ tool can provide promising results for the problems found regarding the usage of high graphical technologies in content creation.

VI. Conclusion and Future Research

LUMOZ is a tool for real time augmented reality application for the digital media live productions. LUMOZ can provide features like live steaming object placement, point tracking and identifications, gestures recognition and 3D data visualization object creations and live 3D visual object placement and 3D object library. All these features are provided for digital content creators who are in live streaming like News productions, live on social medias, live events, online education, and other quick productions.

Currently there are tools and software's for Augmented Reality and 3D creations. There are free tools that anyone can use but with minimum quality and for a quality production there are high budget tools as well. But also, many tools in the current market are for a pre-production. In news productions, sometimes these high graphical works are outsourced to 3rd party companies which can cost a high

budget. LUMOZ can provide content creators a tool with high quality, low production budget and employee budget.

LUMOZ has a wide range of potential market. With the digitalization and current situations in the world, this tool can provide many benefits for both start-up content creators, industrial content creators as well as day to day content creators. Live media broadcasts, live events, and social media, as an online education tool, live News broadcasts and live advertisement productions. As this tool cost less budget and maximum efficiency, the product can focus on high range of target audience. As soon as the implementation process is completed, LUMOZ beginner version will be released to the above identified market potential. As for the future, this research will continue to be carried out so that the product can come up with more updates according to the customer feedbacks.

REFERENCES

[1] Stöber, R. (2004). What Media Evolution Is. European Journal of Communication, 19(4), 483–505. doi:10.1177/0267323104049461

[2] Aitamurto, T., Aymerich-Franch, L., Saldivar, J., Kircos, C., Sadeghi, Y., & Sakshuwong, S. (2020). Examining augmented reality in journalism: Presence, knowledge gain, and perceived visual authenticity. New Media & Society, 146144482095192. doi:10.1177/1461444820951925

[3] Mohamed Zayani (2020): Digital Journalism, Social Media Platforms, and Audience Engagement: The Case of AJ+, Digital Journalism, DOI: 10.1080/21670811.2020.1816140

[4] Scheibe, K., Fietkiewicz, K.J., Stock, W.G.: Information behavior on social live streaming services. J. Inf. Sci. Theory Pract. **4**, 6–20 (2016). https://doi.org/10.1633/jistap.2016.4.2.1CrossRefGoogle Scholar

[5] Y. Genc, S. Riedel, F. Souvannavong, C. Akinlar and N. Navab, "Marker-less tracking for AR: a learning-based approach", Proceedings. International Symposium on Mixed and Augmented Reality. Available: 10.1109/ismar.2002.1115122 [Accessed 24 February 2021].

[6] U. Neumann and S. You, "Natural feature tracking for augmented reality", IEEE Transactions on Multimedia, vol. 1, no. 1, pp. 53-64, 1999. Available: 10.1109/6046.748171 [Accessed 25 February 2021].

[7] Khandelwal, Pulkit et al. "Detection Of Features To Track Objects And Segmentation Using Grabcut For Application In Marker-Less Augmented Reality". Procedia Computer Science, vol 58, 2015, pp. 698-705. Elsevier BV, doi:10.1016/j.procs.2015.08.090. [Accessed 6 Mar 2021].

[8] Park, J., Sung, M. and Noh, S., 2021. Virtual Object Placement in Video for Augmented Reality.

[9] Druzhkov, P., Erukhimov, V., Zolotykh, N., Kozinov, E., Kustikova, V., Meerov, I. and Polovinkin, A., 2021. New object detection features in the OpenCV library.

[10] Dl.acm.org. 2021. An online video placement policy based on bandwidth to space ratio (BSR) | Proceedings of the 1995 ACM SIGMOD international conference on Management of data. [online] Available at: <https://dl.acm.org/doi/abs/10.1145/223784.223853> [Accessed 8 March 2021].

[11] Bacher, I., Javidnia, H., Dev, S., Agrahari, R., Hossari, M., Nicholson, M., Conran, C., Tang, J., Song, P., Corrigan, D. and Pitié, F., 2021. An Advert Creation System for 3D Product Placements. [online] arXiv.org. Available at: <https://arxiv.org/abs/2006.15131> [Accessed 8 March 2021].

[12] Yang, B., Vargas Restrepo, C., Stanley, M. and Marsh, E., 2021. Truncating Bar Graphs Persistently Misleads Viewers.

[13] Qin, X., Luo, Y., Tang, N. and Li, G., 2021. Making data visualization more efficient and effective: a survey.

[14] Luo, W., 2021. User choice of interactive data visualization format: The effects of cognitive style and spatial ability.

[15] Wang, Y., Ren, A., Zhou, M., Wang, W., & Yang, X. (2020). A novel detection and recognition method for continuous hand gesture using FMCW radar. IEEE Access: Practical Innovations, Open Solutions, 8, 167264–167275.

[16] Marathe, V., & Korake, D. (n.d.). *A real time static hand gesture recognition for human computer interaction*. Ezeichen.Com. Retrieved September 7, 2021, from http://www.ezeichen.com/gallery/2-apr-362.pdf

[17] Nagabhyru, S., & Joshi, A. (n.d.). *Gesture Recognition for User Interaction*. Ijsret.Com. Retrieved September 7, 2021, from https://ijsret.com/wp-content/uploads/2021/07/IJSRET_V7_issue4_562.pdf

[18] Almansoury, F., Kpodjedo, S., & Boussaidi, G. E. (2020). Investigating Web3D topics on StackOverflow: a preliminary study of WebGL and Three.js. *The 25th International Conference on 3D Web Technology*.

[19] Wang, D., Chen, S., Li, X., Zhang, W., & Jin, H. (2020). Research on rotational angle measurement for the smart wheel force sensor. *Sensors (Basel, Switzerland)*, *20*(4), 1037.

[20] Hands. (n.d.). Retrieved July 3, 2021, from Github.io website: *Hands*. (n.d.). Github.Io. Retrieved September 7, 2021, from https://google.github.io/mediapipe/solutions/hands

[21] Zhang, F., Bazarevsky, V., Vakunov, A., Tkachenka, A., Sung, G., Chang, C.-L., & Grundmann, M. (2020). MediaPipe Hands: On-device real-time hand tracking. In *arXiv [cs.CV]*. http://arxiv.org/abs/2006.10214

[22] Joy, D. T., Kaur, G., Chugh, A., & Bajaj, S. B. (2021). COMPUTER VISION FOR COLOR DETECTION. *International Journal of Innovative Research in Computer Science & Technology*, *9*(3). https://doi.org/10.21276/ijircst.2021.9.3.9

[23] *StreamYard*. (n.d.). Streamyard.Com. Retrieved September 7, 2021, from *https://streamyard.com/?utm_source=google&utm_medium=cpc&utm_campaign=3-branded&utm_content=119379464307&utm_term=streamyard&gclid=Cj0KCQjw-NaJBhDsARIsAAja6dOeNt5CeaAGv9mFh6UAfaqjV0PGlOJReUk96ce0WM-pm8R0iupD5bQaAnjYEALw_wcB*

[24] *Viz Artist - Broadcast graphic design tool, live motion graphic, virtual sets, augmented reality and 3D animations*. (n.d.). Vizrt.Com. Retrieved September 7, 2021, from https://www.vizrt.com/products/viz-artist

[25] Pavan Kumar, I., Hara Gopal, V. P., Ramasubbareddy, S., Nalluri, S., & Govinda, K. (2020). Dominant color palette extraction by K-means clustering algorithm and reconstruction of image. In *Advances in Intelligent Systems and Computing* (pp. 921–929). Springer Singapore.