Journal of Metaverse
Research Article
Received: 2021-11-16 | **Reviewing:** 2021-11-16 & 2021-12-21 | **Accepted:** 2021-12-21 | **Online:** 2021-12-21 | **Issue Date:** 2021-12-31
Year: 2021**, Volume:** 1**, Issue:**1**, Pages:**9-12

Computer Vision in the Metaverse

Kemal Gökhan NALBANT
Beykent University, Software Engineering
İstanbul, Turkey
kemalnalbant@beykent.edu.tr
0000-0002-5065-2504

Şevval UYANIK
Software Engineering
Beykent University
İstanbul, Turkey
sevvaluyaniik@gmail.com

Abstract **- Metaverse is a rapidly developing new technology today. The purpose of this article is to examine this technology from a computer vision and general perspective. In this study, a comprehensive review of Metaverse concepts in computer vision has been made. Its history, process, techniques, architecture, advantages, and disadvantages are mentioned. The adaptation of Metaverse to life, the ideas of companies about this technological change, and how society will take place in Metaverse are also discussed. The future of Metaverse and what needs to be done to adapt to this technology are explained. As a result, since there are few studies in the literature, this article aims to be a review article that increases academic studies.**

Keywords— ***Metaverse, computer vision, virtual world, extended reality, augmented reality, blockchain, artificial intelligence in metaverse, architecture of the metaverse.***

I. INTRODUCTION

Since the term metaverse is a newly emerging term, it has been observed that there are not many studies in the field of computer vision when the literature is searched. Therefore, this article is a review study. Metaverse, the combination of the prefix "meta" (imply-ing transcending) with the word "universe," describes a hypothetical synthetic environment linked to the physical world. The term 'metaverse' was first coined in a speculative novel named Snow Crash, written by Neal Stephenson in 1992 [1]. In this novel, Neal Stephenson described it as a 3D virtual world where people exist as avatars and interact with software agents. The term Digital Twins was first used by Michael Grieves at the Society of Manufacturing Engineers conference at the University of Michigan in 2002. Digital Twins was introduced as a digital equivalent of a physical object and the conceptual model underlying product lifecycle management. Second Life is an online virtual world game developed with Philip Rosedale and his team. In this game, as in Metaverse, human beings exist as avatars and are in a virtual world. This game is at an important point in the Metaverse world. Second Life started to be developed in 2003, and today, it has millions of active users.

Before examining the subject of computer vision in the Metaverse, it will be useful to examine some terms. Subjects to be examined are Extended Reality (XR), Mixed Reality (MR), Augmented Reality (AR) and Virtual Reality (VR).

Extended Reality (XR) refers to real and virtual spaces created by wearable devices. In other words, it can also be called human-machine interaction [2]. Mixed Reality (MR) combines virtual and real worlds to produce new environments in which digital and physical objects interact in real-time. Mixed Reality (MR) does not coexist two realities; it is a hybrid of two realities [3]. Augmented Reality (AR) is a technology used to increase the visual fields of users with necessary information [4]. In Virtual Reality (VR), users are entirely virtual. They have no connection to the real world. Virtual Reality (VR) allows the user to transition into a three-dimensional (3D) virtual world with a computer [5].

Continuing from the history and development of computer vision, described as the father of computer vision, Larry Roberts, while doing his Ph.D. at MIT, first discussed the possibility of extracting 3D geometric information from 2D perspective views in his thesis. In this way, the foundations of computer vision began to form in 1960 [6]. David Marr defined vision as proceeding from a two-dimensional vision to a three-dimensional visual recognition in 1982 [7]. David Marr's approach is one of the most influential studies ever done. David Marr approach uses low-level image processing algorithms applied to 2D images to obtain a 2.5D sketch of the scene and calls it the first draft. It then obtains 3D model representations of the scenes using high-level techniques. After this approach, many researchers supported David Marr. The Marr approach is complicated to execute, and also, in most computer vision applications, 3D models are not required to be imported. Only one purpose-built computer vision application can be used. For example, only face detection may be necessary, not requiring a full 3D model. This new approach defines those algorithms that should be targeted only and are called "Purposeful Vision". The main proponent of this approach is Yiannis Aloimonos of the University of Maryland [8].

II. ADVANTAGES OF METAVERSE IN COMPUTER VISION

Computer vision plays an essential role in building humans' ability to experience the virtual world in the Metaverse universe. Digital avatars in computer vision and virtual reality offer an almost real-world equivalent experience. While you can reach this virtual world using extended reality products, extended reality is based on the basis of computer vision. Computer vision plays a vital role in XR applications. Visual information and computer vision play a vital role in processing, analyzing, and understanding visuals as digital images or videos to derive meaningful decisions and take actions. Computer vision allows XR devices to recognize and understand visual information of users' activities and physical surroundings, helping build more reliable and accurate virtual and augmented environments. Human pose tracking refers to the computer vision task of obtaining spatial information concerning human bodies in an interactive environment [9].

Computer vision, in XR applications, recreates the user environment in 3D. It determines the direction and location of the user. XR interactive system needs to track the body and pose of users. Metaverse, the human users will be followed with computer vision and represented as avatars. Metaverse also requires users to perceive their environment. Therefore,

image processing is an important area for developing a better metaverse. The metaverse is connected seamlessly with the physical environments in real-time. An avatar needs to work with a physical person in such a condition. It is crucial to display the 3D virtual world with less noise, blur, and high resolution in the metaverse. In adverse visual requirements, such as haze, low or high luminosity, or even rainy weather conditions, the interactive systems in the metaverse still need to show the virtual universe [10].

III. Disavantages of Metaverse in Computer Vision

One of the significant disadvantages of computer vision on Metaverse is the inability to access a fast and secure internet resource and the difficulty in making the relevant extended reality (XR) technology accessible to everyone [11]. Since a vast virtual universe is mentioned, it may not be accessible in the future to make these enhancements in this sizeable virtual universe and keep the user experience always at maximum performance when using computer vision. Also, as the number of users begins to increase, these technologies will always need to work up to date, smoothly, and at high performance. These technologies will always need to be developed.

IV. Applications of Metaverse in Computer Vision

Metaverse domains include healthcare, military, real estate, manufacturing, education, and retail. Metaverse Health applications accelerate surgical procedures, examine data obtained by 3D scans in real-time and augmented reality, help patients detect, diagnose and treat possible diseases, and allow patients to look at their bodies.

Some Metaverse military applications provide a more realistic experience by putting soldiers in more physically and psychologically challenging war environments.

A. Metaverse Real Estate Applications

It allows real estate marketers to review the property with virtual reality tours before making any selections for their clients. In this way, it offers customers the advantage of inspecting under all conditions and reduces travel time to zero.

B. Metaverse Production Applications

Factories significantly reduce the likelihood of accidents by using virtual reality applications. You also do not need to spend time and money training a new employee because, with augmented reality simulations, employees can be prepared quickly and safely.

C. Metaverse Education Applications

Traditional teaching approaches can never provide the high efficiency of using concepts visually. Students prefer watching rather than reading. For this reason, dealing with a subject with visual dimensions with virtual reality will always give better learning results.

D. Metaverse Retail Applications

One of Metaverse's most significant recent developments is that Meta (formerly Facebook) plans to launch direct-to-consumer virtual stores with virtual reality and augmented reality. This promising news also plays an encouraging role in charting different routes in Metaverse.

V. Computer Vision Techniques

A. Image Classification

Classifying pixel and vector groups in an image by applying specific rules is called image classification. It is among the best-known techniques. However, it has several problems to be overcome in its implementation. Let's say we have a set of images belonging to a category, and we prepared a test image set to measure the accuracy of our predictions. The difficulties here are; changing perspectives, deformation, and light settings are some points that must be overcome.

B. Object Detection

Object detection is a method used to identify and locate specific objects in any image or video. Object detection allows us to determine the positions or movements of said objects in the scene and draws them with the bounding box. The major difference between object detection and image recognition is that it defines and labels a bounding box for every object in an image or video defined as a specific object. The model predicts the object's label and location.

C. Object Tracking

Object tracking is defined as tracking a moving object in any or many scenes. First, object detection is used, and then deep learning applications are used, where the object's movement is monitored. Tracked objects have an indicator next to them. The bounding box that shows where the object is can be seen in Fig.1.

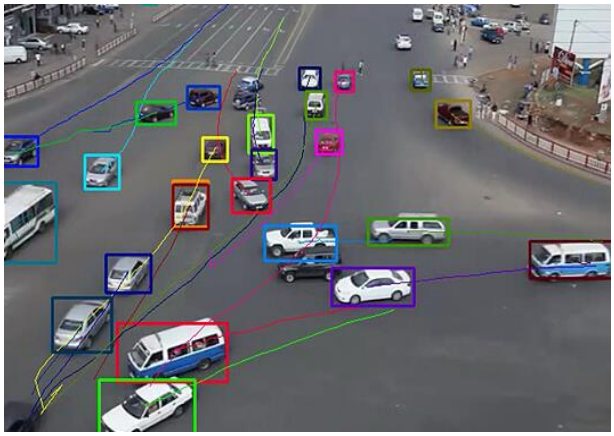

Fig. 1. Bounding Box in Object Tracking [12]

VI. Impact of Covid-19 on Metaverse

Before Covid-19, Metaverse was seen as just an entertainment platform for the community, like spending time, playing games. But after the pandemic, a potential second world came to the Metaverse for society. The community is discussing business integration, collaborations, retail merchandising, investment, and how to adapt this model to life around the world. Building a virtual universe has gained extreme importance after the pandemic, as it can be a solution to some problems. An example of this is holding interactive virtual business meetings. The Covid-19 pandemic has played an important role in accelerating the evolution of the Metaverse universe.

VII. Future of Metaverse in Computer Vision

It looks like Metaverse will be at an important point in our lives in the future. Therefore, many investments are made. Today, while the business world is trying to adapt quickly, it is predicted that many sectors will take place in the Metaverse in the future. From the point of computer vision, extended reality (XR) and augmented reality (AR) technology products will develop, or new products will begin to emerge in the future.

VIII. Blockchain on Metaverse

Blockchain technology is originally the name given to the design underpinning the operation of the digital currency Bitcoin. Bitcoin's creator never used the term "blockchain" in his whitepaper. Reading the paper, one gets the distinct impression that the author was not introducing a new technology in the traditional sense of the word. Still, a software design drawing on several existing technologies allows him to create a "purely peer-to-peer version of electronic cash" [13]. There are some projects in Metaverse. Metaverse also aims to create a digital economy. It has cryptocurrencies for it. The benefits of Blockchain in the Metaverse are as follows; digital proof of ownership, digital collectability, transfer of value, governance, accessibility, and interoperability. It also offers a transparent and cost-effective solution.

IX. Artificial Intelligence In Metaverse

Artificial intelligence applications; Features deep learning provide higher performance for developers and designers in Metaverse than traditional approaches. However, it is not enough to implement artificial intelligence to make users' work easier and provide a high-performance experience. Existing AI models are often intense and require large computations. Therefore, it is essential to design light and efficient artificial intelligence models [14].

X. Extended Reality (xR)

The Metaverse moves from concept to reality and requires a VR/AR/MR intermediate stage. To some extent, virtual environments form the technical basis. Technologies like Metaverse will shape the new form of the internet. VR will also make the virtual world operation more similar to the real world, allowing users to have a more realistic and memorable experience. On the other hand, AR/MR can transform the physical world into a virtual one. The future of the physical world is being integrated into the Metaverse.

Ideal model; MR and Metaverse advocate full integration of virtual entities with the physical world. Thanks to these technologies, users with digital assets can interact and collaborate with objects [15]. Also, some keys are given for the Metaverse in Fig.2.

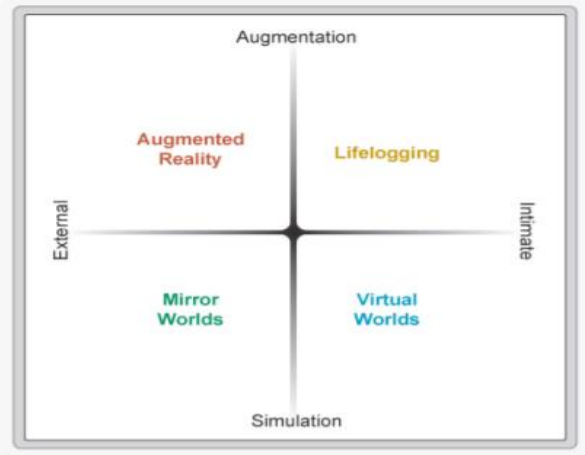

Fig. 2. Metaverse Keys [16]

A. Augementation

It refers to technologies that add new features to the existing system. In Metaverse, it means technology that adds new information to our perception of the environment.

B. Simulation

It refers to technologies that model reality and offer new environments. In Metaverse, it means technologies that provide simulated worlds in the Metaverse.

C. Intimate Technologies

It refers to the identity and actions of users or objects. In Metaverse, it means technologies in which the user or object is the avatar or actor representative in the system.

D. External Technologies

It refers to the outside, that is, to the general world. In Metaverse, it means technologies that control the world around the user.

XI. Architecture Of The Metaverse

The Metaverse is still in the active phase of development. Therefore, there is no consistent definition of architecture. Jon Radoff proposed a seven-layered architecture from bottom to top [17]. These; infrastructure, human interface, decentralization, spatial computing, creator economy, discovery, and experience. This architecture represents the industrial division. However, the Metaverse architecture should be viewed from a more micro perspective. A three-layer architecture is shown in Fig.3. These; infrastructure, interaction, and ecosystem. The basic requirements of the Metaverse require the transition of the architecture from a physical world to a virtual one. These are briefly mentioned in Fig 3.

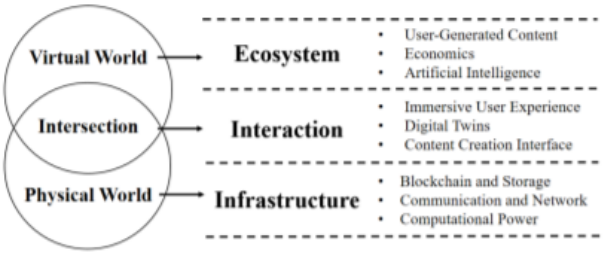

Fig. 3. Three-layer Architecture of the Metaverse [18]

XII. Conclusions

In conclusion, Metaverse is an essential and promising topic today. It is foreseen that the Metaverse approach will focus on how this technology will be adapted to people's lives rather than how this technology will be developed in the future. For this reason, this technology currently needs to be further researched and developed. The subjects that need to be developed are computer vision, augmented reality, and extended reality technologies, which will make a faster and positive contribution to Metaverse. Companies need to invest more in this direction, and technologies emerge quicker and more advanced. Especially in the business world, the sector's leading companies should take place in Metaverse and set an example in this regard. This vision's rapid acquisition and realization play an essential role in our future. In addition, these technologies need to be more accessible in society. This technology, which will be an integral part of our lives soon, should be suitable for people at every economic level of the organization. Instead of a high-cost product, cheaper-cost products should be developed and made fit for the use of all people.

References

[1] Joshua, J. (2017). Information Bodies: Computational Anxiety in Neal Stephenson's Snow Crash. Interdisciplinary Literary Studies, 19(1), 17-47.

[2] North of 41 (2020). What really is the difference between AR / MR / VR / XR? Access Date: 15/12/2021 https://medium.com/@northof41/what-really-is-the-difference-between-ar-mr-vr-xr-35bed1da1a4e

[3] Milgram, Paul & Kishino, Fumio. (1994). A Taxonomy of Mixed Reality Visual Displays. IEICE Trans. Information Systems. vol. E77-D, no. 12. 1321-1329.

[4] Thomas, P. C., & David, W. M. (1992, January). Augmented reality: An application of heads-up display technology to manual manufacturing processes. In Hawaii international conference on system sciences (pp. 659-669).

[5] Fast-Berglund, Å., Gong, L., & Li, D. (2018). Testing and validating Extended Reality (xR) technologies in manufacturing. Procedia Manufacturing, 25, 31-38.

[6] Huang, T. S. (1996). Computer vision: Evolution and promise. CERN European Organization for Nuclear Research-Reports-CERN, 21-26.

[7] Marr, D. (1982). Vision: A computational investigation into the human representation and processing of visual information, henry holt and co. Inc., New York, NY, 2(4.2).

[8] Huang, T. S. (1996). Computer vision: Evolution and promise. CERN European Organization for Nuclear Research-Reports-CERN, 21-26.

[9] Barioni, R. R., Figueiredo, L., Cunha, K., & Teichrieb, V. (2018, October). Human Pose Tracking from RGB Inputs. In 2018 20th Symposium on Virtual and Augmented Reality (SVR) (pp. 176-182). IEEE.

[10] Fast-Berglund, Å., Gong, L., & Li, D. (2018). Testing and validating Extended Reality (xR) technologies in manufacturing. Procedia Manufacturing, 25, 31-38.

[11] Marion Davies (2021). Pros and Cons of the Metaverse. Access Date: 15/12/2021. https://www.konsyse.com/articles/pros-and-cons-of-the-metaverse/

[12] Mehul (2020). Object Tracking in Videos: Introduction and Common Techniques. Access Date:15/12/2021. https://aidetic.in/blog/2020/10/05/object-tracking-in-videos-introduction-and-common-techniques/

[13] Ammous, S. (2016). Blockchain Technology: What is it good for?, Available at SSRN 2832751.

[14] Lee, L. H., Braud, T., Zhou, P., Wang, L., Xu, D., Lin, Z., ... & Hui, P. (2021). All one needs to know about metaverse: A complete survey on technological singularity, virtual ecosystem, and research agenda. arXiv preprint arXiv:2110.05352.

[15] Lee, L. H., Braud, T., Zhou, P., Wang, L., Xu, D., Lin, Z., ... & Hui, P. (2021). All one needs to know about metaverse: A complete survey on technological singularity, virtual ecosystem, and research agenda. arXiv preprint arXiv:2110.05352.

[16] Smart, J., Cascio, J., & Paffendorf, J. (2007). Metaverse Roadmap 2007: pathways to the 3D Web. A Cross-industry Public Foresight Project. Retrieved December, 31, 2008.

[17] Jon Radoff (2021). The Metaverse Value-Chain. Access Date: 15/12/2021. https://medium.com/building-the-metaverse/the-metaverse-value-chain-afcf9e09e3a7

[18] Duan, H., Li, J., Fan, S., Lin, Z., Wu, X., & Cai, W. (2021, October). Metaverse for social good: A university campus prototype. In Proceedings of the 29th ACM International Conference on Multimedia (pp. 153-161).