NONINTRUSIVE VIRTUAL TRAINER PROTOTYPE FOR EXERCISE ROUTINES IN UNITY USING MOTION CAPTURE

Prototipo de entrenador virtual no intrusivo para rutinas de ejercicio en unity utilizando captura de movimiento

GERARDO JOSÉ MORENO URRIOLA[1], MARTIN JOSE VEGA BONILLA[2], CRISTIAN IVÁN PINZÓN TREJOS[3]

Recibido:30 de noviembre de 2022. Aceptado:16 de enero de 2023

DOI: http://dx.doi.org/10.21017/rimci.2023.v10.n19.a125

Abstract

Virtuality is an innovative process seen in recent years. This process allows transforming scenarios to a virtual environment for simulations, practices, or tests so that all results can be studied. In turn, these environments can be supported by new tools and methodologies that increase the analysis capabilities, such as motion capture to generate animations, tracking of robotic equipment and the study of human movement, which benefits the development of systems with these objectives. A prototype of a nonintrusive virtual trainer capable of capturing motion and determining the correct execution of exercise routines is presented, using a virtual environment developed in the Unity video game engine. The system employs a SHDR webcam for real-time capture of the movement performed by the user, which is processed to track the pose and joints using Machine Learning through the MediaPipe library. This article explains the construction of the prototype and presents the results of the project.

Keywords: Motion capture; virtual environments; Machine Learning; exercise routines; Unity.

Resumen

La virtualidad es un proceso innovador visto en los últimos años. Este proceso permite transformar escenarios a un entorno virtual para realizar simulaciones, prácticas o pruebas de manera que se puedan estudiar todos los resultados. A su vez, estos entornos pueden apoyarse en nuevas herramientas y metodologías que aumentan las capacidades de análisis, como la captura de movimiento para generar animaciones, el seguimiento de equipos robóticos y el estudio del movimiento humano, lo que beneficia el desarrollo de sistemas con estos objetivos. Se presenta un prototipo de entrenador virtual no intrusivo capaz de capturar el movimiento y determinar la correcta ejecución de las rutinas de ejercicio, utilizando un entorno virtual desarrollado en el motor de videojuegos Unity. El sistema emplea una cámara web SHDR para la captura en tiempo real del movimiento realizado por el usuario, que se procesa para rastrear la pose y las articulaciones mediante Machine Learning a través de la biblioteca MediaPipe. Este artículo explica la construcción del prototipo y presenta los resultados del proyecto.

Palabras clave: Captura de movimiento; entornos virtuales; aprendizaje automático; rutinas de ejercicio; Unity.

1 Bachelor's Degree in Computer Systems Engineering Universidad Tecnológica de Panamá. Veraguas, Panamá. ORCID: https://orcid.org/0000-0001-9233-8815 Correo electrónico: gerardo.moreno2@utp.ac.pa

2 Bachelor's Degree in Computer Systems Engineering Universidad Tecnológica de Panamá. Veraguas, Panamá. ORCID: https://orcid.org/0000-0002-1647-8928 Correo electrónico: martin.vega @utp.ac.pa

3 School of Computer Systems Engineering Universidad Tecnológica de Panamá. Veraguas, Panama. ORCID: https://orcid.org/0000-0003-0846-5366 Correo electrónico: cristian.pinzon@utp.ac.pa

I. Introduction

Virtual Environments (VE) have had a growing impact in recent years thanks to the adaptability they have in simulating any type of real scenario that has a type of interaction with the user. There are applications focused on medical, mechanical, and even spatial environments, where the user or patient can achieve an experience almost equal to reality.

The properties of the VE to be developed depend on the task to be solved. Generally, these environments are accompanied by tools and technologies that provide a certain degree of interactivity with respect to an effect. Studies such as[1] demonstrate the effectiveness of VEs together with electromyographic (EMG) sensor equipment to anticipate and estimate the position of the head, which improves the sensation of movement for Head-Mounted Displays (HMD) in simulators.

On the other hand, applications are seen as[2] and[3] focused on clinical medicine and neuropsychology, respectively, where cognitive rehabilitation methodologies and visuospatial medical assessments are applied for patient treatment. By including processing such as virtual reality (VR), refined VEs are obtained in the study and assessment of a user's functional capabilities.

Regarding the specific case of this study, motion capture is another new technology integrated as a tool to be applied to VEs, thanks to its adaptability and real-time model reconstruction using reflective markers. The literature review in[4] demonstrates the growing influence of this equipment for its application in human motion analysis, indicating a general efficiency in those systems validated to be applied in environments with biomechanical studies.

However, there are nonintrusive systems seen in[5] that apply motion capture using a processing known as volume intersection which solves, to a large extent, the use of specialized motion capture hardware to accomplish this task. Image processing coupled with the development of convolutional neural networks (CNNs) is another nonintrusive method for the application of motion capture on less expensive equipment[6].

This reduces the need to acquire physical equipment focused on motion capture; thanks to this means, new VEs have been developed that use a simple webcam for motion detection, so it is possible to study and analyze the motion performed by an object or a user remotely or at any workstation.

The application of these systems benefits those situations where there is no access to a specialized center, either due to high costs or mobility restrictions, as seen in recent years due to global epidemics. This problem is visualized in the impediment that some individuals have in exercising these practices in Panama. An example of this was the adaptation of educational centers in subjects such as physical education to cope with virtual practice[7]. One solution lies in VEs, together with tools such as motion capture, which have made it possible to facilitate the experience in environments of different areas, both industrial and professional processes or daily activities, dance centers, training, and rehabilitation[8][9][10].

From this problem, the following question arises: What is the appropriate architecture to build a virtual environment that can perform motion capture in a nonintrusive way? A functional, inexpensive, and nonintrusive system can become a useful tool for those individuals who cannot experience these activities at home.

In view of the approach, this paper aims to present the methods and implementations necessary to develop a prototype VE for the simulation of a virtual trainer using motion capture technologies. The system bases its programming on daily exercise routines so that the user can simulate them by pose tracking and experience training in a real gym from a VE designed in Unity.

Finally, the article is structured as follows: The first section introduces the reader about the issues focused in the research study, then the second section explains the methodology applied to the study. The third section indicates the tools and methods for the development of the prototype, as well as the data collection and testing. The fourth section analyzes the results obtained and finally, the fifth section presents the conclusions of the project.

II. Methodologies

A. Research methodology

The following figure shows the steps of the methodology used.

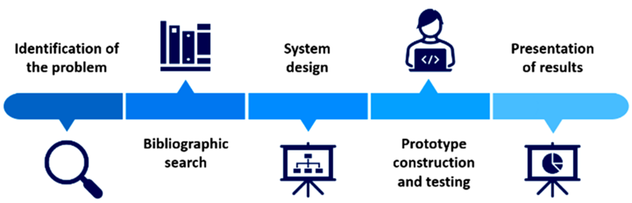

Fig. 1. Steps of the applied research methodology.

B. Software methodologies

The prototype methodology was used for software development, starting with the first step to know the specifications of the system. After this, the conceptual design and creation of the prototype was developed, and the system tests were started in order to evaluate the effectiveness of the software used. Once the tests are completed, the results are evaluated to apply the identified improvements to the prototype to obtain the desired system.

Fig. 2. Cycle for the development of software prototypes.

III. Materials and methods

The development of this prototype consisted of the implementation of both hardware and software equipment. Each of them will be explained.

A. Technologies involved

Both hardware and software equipment and components were used to build the prototype. Each of these is described below.

HD Webcam SHDR

The SHDR (Super High Dynamic Range) Webcam[11] is a camera integrated into a computer with the ability to capture images of different levels of light exposure, which overexposes the captured textures improving the graphic quality.

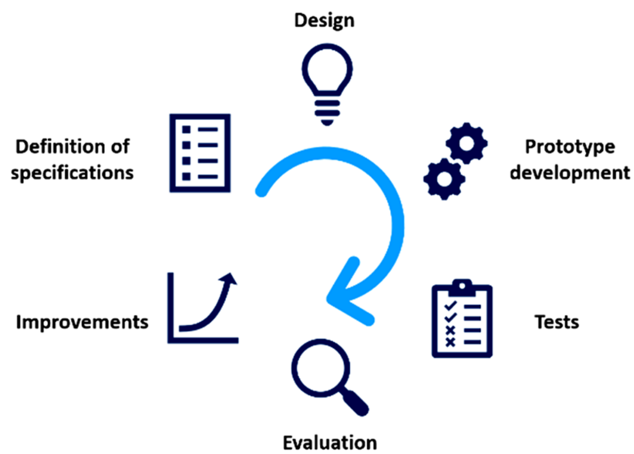

Fig. 3. SHDR webcam.

B. Software used

The following software tools and environments were used to manage information processing.

Python 3.10

Python 3.10[12] is a multipurpose object-oriented programming language, versatile for the development of artificial intelligence environments and programs, Machine learning, Deep learning, in addition to the study of Data Science and web development.

Unity

Unity 3D[13] developed by Unity Technologies, is an engine or tool for the development of video games and multiplatform virtual environments. It integrates an event interpreter to design scenarios, animations, image rendering and provides a physics engine that allows giving effects to the actions that occur in an environment.

Blender

The Blender suite[14] is a tool for the development and design of 3D or three-dimensional models integrating various platforms for graphic design. It provides rendering, video editing, texturing and simulation methods, making it a professional platform for computer media production.

Libraries

For the development of image processing and pose tracking, the following libraries were used:

- OpenCV[15] // For the integration of vision and image reading tools.
- MediaPipe[16] // For the use of modules with Machine Learning utilities together with real-time audio and video processing.
- Socket[17] // For running nodes on a local network so that they can communicate via an IP and a computer port.

C. Participants

We proceeded to conduct a study on the behavior of the prototype, using a convenience sampling by selecting a set of 6 participants, with or without experience performing daily exercise routines. A table with the participants' information is shown below:

Table I. Participants selected for the tests

Num.	Age	Genre	Weight (Kg)	Height (cm)
1	22	M	67	178
2	22	M	70	173
3	21	M	64	166
4	22	M	94	171
5	22	M	65	182
6	21	M	100	184

D. Data collection

For data collection in the tests performed, each participant was instructed to perform a set of 10 tests, or repetitions, in the four modalities of routines integrated into the system: right elbow flexion, left elbow flexion, squats and scissor jumping. For each of the tests, the total time required by the participant to complete the sessions was taken.

Fig. 4. Participant performing the tests together with the VE during data collection.

The tests were captured in real time showing the operation of the prototype to the participants while they simulated the movement indicated by the system. Once they have completed the routine, they take a two-minute break to continue with the next exercise.

E. System execution

The prototype starts by running the Python system for capturing images in real time by means of the SHDR webcam on the computer used. The camera takes the first capture applying a filter that creates a segmentation of the human body so that the pose is estimated by a DataSet processed via machine learning from the interpreter; this generates the reference points in the pose tracking in the image. Once the system estimates the poses and the location of the body joints with respect to the depth level, the points are stored in a list that is sent to the VE in Unity through a Socket with the IP and port indicated in the equipment. Upon receiving this data, the VE hosts the locations of the points to a set of traces that assimilate the image captured from Python, which simulates the captured motion.

Then, using programming interpreters in the Unity VE, the vector positions of the limbs are obtained together with the pose of the upper and lower part of the human body. This generates a desired angle thanks to the vector values, which is compared with a standardized angle of the pose to be performed by the participant to identify whether or not it is performing the exercise correctly, so that it is shown in the VE for visual effects.

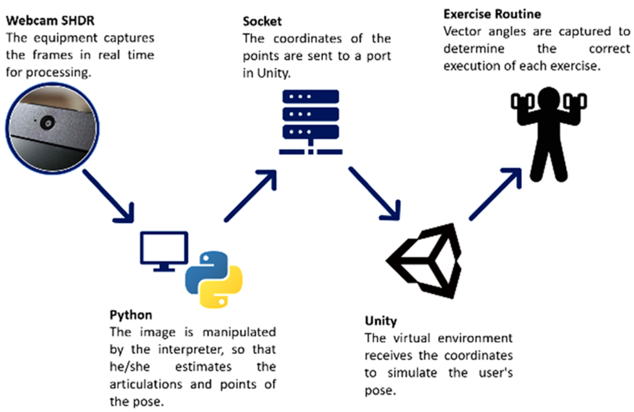

Fig. 5. Conceptual design of the developed prototype.

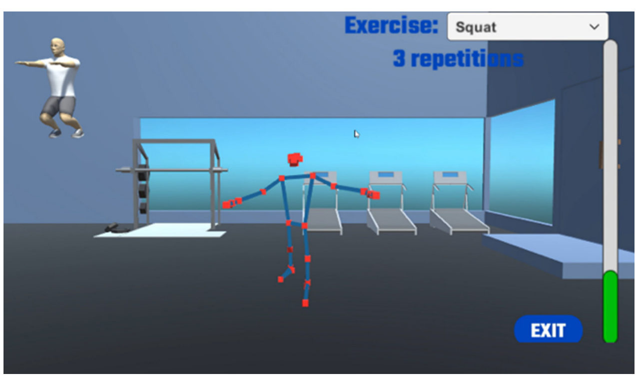

Fig. 6. Unity virtual environment simulating the captured motion.

IV. Results

A survey focused on knowledge and interest in the tool was conducted. A total of 32 surveys were applied. According to the data obtained, 88% of the people surveyed have come to perform physical exercises before, while 12% have not come to perform a type of exercise safely.

As for users who have gone to a gym, 50% of them have attended one of these centers to exercise. Finally, 94% of the participants indicated that they would use the tool to perform exercise routines with support.

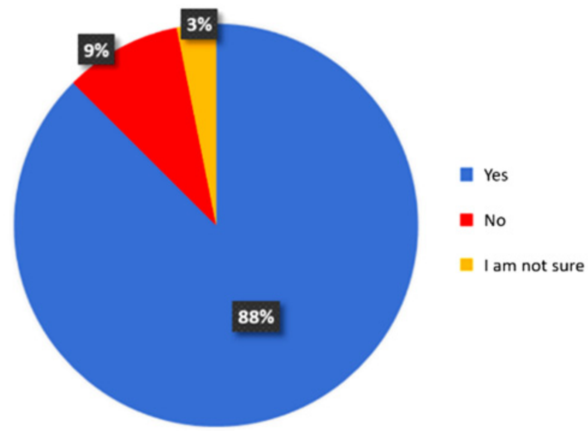

Fig. 7. Survey section: Have you done physical exercise before?

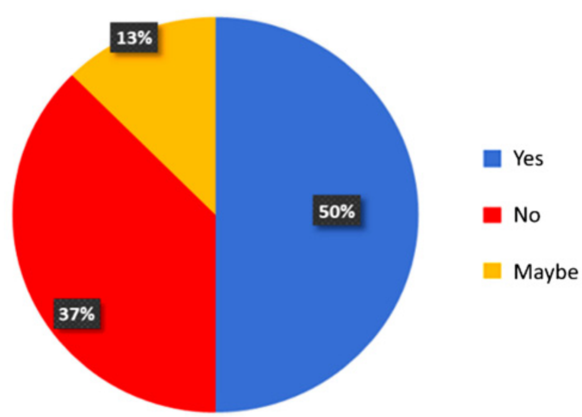

Fig. 8. Survey section: Have you ever needed to attend a gym or specialized center for exercise before?

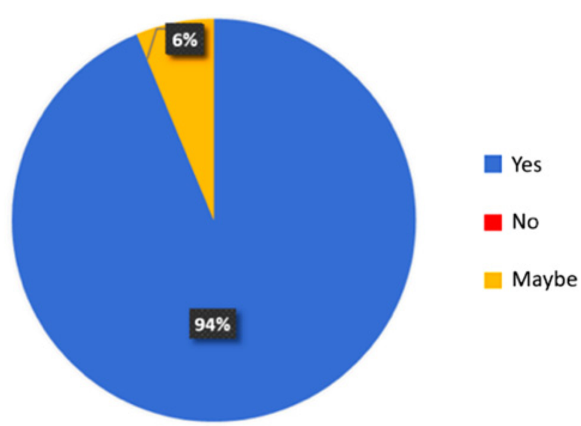

Fig. 9. Survey section: Would you agree to use a virtual environment that captures your movements to show and tell you whether you are performing an exercise routine correctly?

Participants were given 10 tests for each of the exercise routines, or modalities, so that a total of 60 tests were captured for each modality. Next, the tests will be analyzed depending on the system's successes and failures in each routine:

A. Tests for right elbow flexion

This modality registered a total of 58 successes, which reflects a 96.67% effectiveness with the prototype. The results obtained demonstrate the degree of efficiency that the system presents for the solution of the indicated problem, in addition to its capacity to correctly determine the angular value of the movement performed by the participants.

The tests identified as failures depended on the correct detection of each session in the exercise routines, since these were performed continuously, even though the system does not detect or classify the execution of the movement in succession. In this way, it is possible to analyze the degree of precision that the prototype possesses when faced with different angles and body proportions.

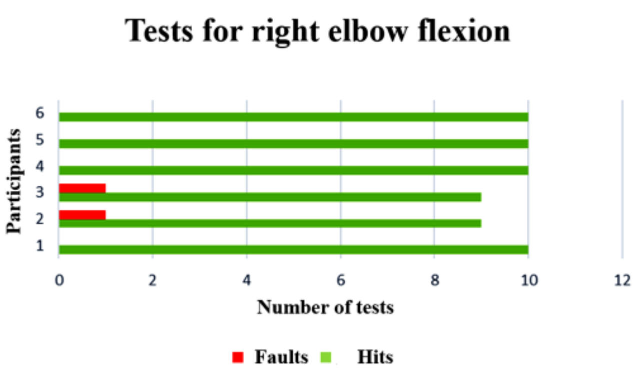

Fig. 10. Test results for right elbow flexion.

B. Left elbow flexion tests

Fig. 11. shows the results obtained for the exercise routines modality in left elbow flexion. A greater detection of failures is highlighted in contrast to the previous modality, obtaining 70.0% effectiveness.

Several reasons why the prototype was more prone to miss the participants' arm position were analyzed. One possible reason is incorrect placement of the participant with respect to the desired pose, so that the depth generated by the interpreter produced a value that was not expected.

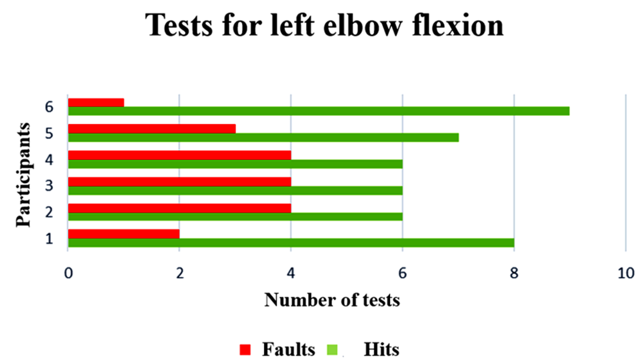

Fig. 11. Test results for left elbow flexion.

On the other hand, there is the possibility that the lighting and contrast of the image resulted in incorrect tracking, which generated failures to identify whether the routine was executed correctly.

C. Squat tests

For this modality, the results reflected positive values. A 95.0% effectiveness rate was obtained with 57 hits and 3 misses, so that the prototype identified the sequences correctly even with the misses captured in previous modalities.

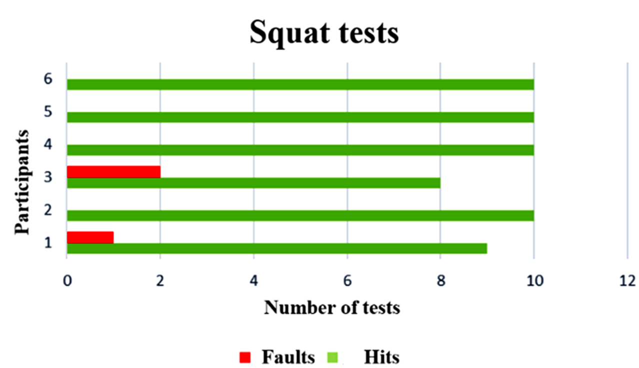

Fig. 12. Test results for squats.

D. Scissor Jumping Tests

The last modality performed by the participants was the routine with the best results. Fig. 13 reflects 98.33% effectiveness with only one failure identified, which is a crucial result to visualize the methodologies and applications that the system has.

It is necessary to emphasize the fact that the participants were in constant motion for this modality, in addition to the variations with respect to the pose tracking obtained by the prototype together with the VE. Sequentially, the speed of image processing was sufficient to show a clear picture of the movement performed by the participants.

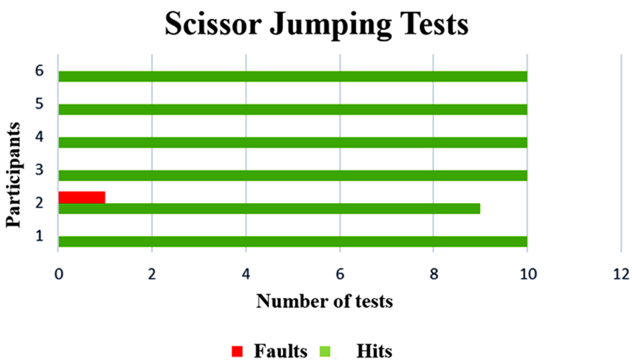

Fig. 13. Test results for scissor jumps.

The technologies integrated into the project demonstrate a high degree of efficiency by acting as a method to support the participants in performing the movements correctly. This has a positive influence on the validation of the system, so that it can be used in different environments. On the other hand, the participants showed interest in the tool thanks to its simplicity and easy execution. This is because the exercise routines were developed in such a way that any individual can perform them.

The performance of the system solves to a great extent the problem of attending a conventional center to exercise these practices, in addition to the capacity it has to adapt to various equipment as long as they have a digital camera, being one of the current barriers in the commercialization and application of these tools.

V. Conclusions

This paper has presented the design and development of a prototype of a nonintrusive virtual trainer to support the execution of exercise routines by means of motion capture in a virtual environment in Unity. The main achievements were the following:

A survey was developed and published to determine people's awareness of virtual environments, their applications, and their interest in a virtual trainer for exercise through motion capture.

The various methods and components, both hardware and software, were identified for the precise execution and operation of the prototype, integrating a low-cost motion capture system.

Interfaces were designed to provide the user with a means to visualize the operation and processes managed by the prototype, facilitating the use of the tool regardless of the technological knowledge of the participant.

The prototype developed has a degree of accuracy of 90.0%, taking as a reference the values obtained in the tests carried out in the different modalities of the system.

The ability of motion capture methodologies to be implemented in virtual environments to solve different problems applicable to a wide variety of projects was demonstrated.

It is planned to continue with the construction and the complete development of the project, analyzing new components that allow to generate a higher level of precision than the one obtained in the execution of this research. It is expected to have an innovative application in the future as a marketable product.

References

[1] Y. Barniv, M. Aguilar, and E. Hasanbelliu, "Using EMG to anticipate head motion for virtual-environment applications," IEEE Trans. Biomed. Eng., vol. 52, no. 6, pp. 1078–1093, 2005, doi: 10.1109/TBME.2005.848378.

[2] L. Li et al., "Application of virtual reality technology in clinical medicine.," Am. J. Transl. Res., vol. 9, no. 9, pp. 3867–3880, 2017. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5622235/.

[3] A. Rizzo et al., "Virtual environment applications in clinical neuropsychology," in Proceedings IEEE Virtual Reality 2000 (Cat. No.00CB37048), 2000, pp. 63–70. doi: 10.1109/VR.2000.840364.

[4] G. Nagymate and R. Kiss, "Application of OptiTrack motion capture systems in human movement

analysis A systematic literature review," Recent Innov. Mechatronics, vol. 5, 2018, doi: 10.17667/riim.2018.1/13.

[5] A. Bottino and A. Laurentini, "Experimenting with nonintrusive motion capture in a virtual environment," Vis. Comput., vol. 17, pp. 14–29, 2001, doi: 10.1007/s003710000091.

[6] D. Osokin, "Real-time 2D Multi-Person Pose Estimation on CPU: Lightweight OpenPose." arXiv, 2018. doi: 10.48550/ARXIV.1811.12004.

[7] K. E. Lara, "El doble desafío de la Educación Física virtual," Panamá América: Salud y bienestar, Panamá, Sep. 07, 2020.[Online]. Available: https://www.panamaamerica.com.pa/deportes/el-doble-desafio-de-la-educacion-fisica-virtual-1171545

[8] L. P. Berg and J. M. Vance, "Industry use of virtual reality in product design and manufacturing: a survey," Virtual Real., vol. 21, no. 1, pp. 1–17, 2017. https://link.springer.com/article/10.1007/s10055-016-0293-9

[9] W. A. IJsselsteijn, Y. A. W. de Kort, J. Westerink, M. de Jager, and R. Bonants, "Virtual Fitness: Stimulating Exercise Behavior through Media Technology," Presence Teleoperators Virtual Environ., vol. 15, no. 6, pp. 688–698, Dec. 2006.

[10] J. C. P. Chan, H. Leung, J. K. T. Tang, and T. Komura, "A Virtual Reality Dance Training System Using Motion Capture Technology," IEEE Trans. Learn. Technol., vol. 4, no. 2, pp. 187–195, 2011, doi: 10.1109/TLT.2010.27.

[11] T. Hayami, M. Tanaka, M. Okutomi, T. Shibata, and S. Senda, "Super-high Dynamic Range Imaging," in 2014 22nd International Conference on Pattern Recognition, 2014, pp. 720–725. doi: 10.1109/ICPR.2014.134.

[12] P. Galindo Salgado, "What's New In Python 3.10," 2022. https://docs.python.org/3/whatsnew/3.10.html

[13] Unity Technologies, "Unity," 2022,[Online]. Available: https://unity.com/

[14] Blender Foundation, "Blender: Introducing blender 3.1," 2022. https://www.blender.org/

[15] G. Bradski, "The OpenCV Library" Dr. Dobb's J. Softw. Tools, 2000.

[16] C. Lugaresi et al., "MediaPipe: A Framework for Building Perception Pipelines." arXiv, 2019, https://arxiv.org/abs/1906.08172.

[17] Python, "Socket - Low-level networking interface," 2022. https://docs.python.org/3/library/socket.html