Real-time Hand Tracking Using Leap Motion Controller

Roumaissa Bekiri[1], Mohamed Chaouki Babahenini[1]

[1]*LESIA Laboratory, Computer Science Department, University of Biskra, Algeria*

Abstract

Hand tracking has become a crucial topic in recent years due to the advancement in various applications of human-computer interaction (HCI). The primary objective of this research is to investigate the capabilities and limitations of the Leap Motion Controller in real-time hand tracking scenarios. We begin by providing an overview of the Leap Motion technology, its hardware architecture, and the software stack used for hand tracking. We delve into the underlying algorithms and methods employed by the controller to analyze and interpret hand and finger movements. Then, we propose several scenes to interact with it real-hand through leap motion camera. Our experimental results demonstrate that our approach, relying on tracking data from the Leap Motion Controller, accurately identifies hand gestures in the absence of any occlusion.

Keywords

Leap Motion, hand tracking, gesture, real-time interaction

1. Introduction

Within the realm of gesture recognition systems, tracking stands out as a pivotal component. Over the course of the past year, there has been a noteworthy surge in the development of various techniques and methodologies aimed at enhancing the overall performance of gesture recognition systems. These efforts have predominantly focused on four key aspects: segmentation, tracking, modeling, and recognition'. Tracking is an essential element in the sequence as it involves monitoring and following the movements of the segmented gestures over time. Achieving robust and accurate tracking is vital to ensure the continuity and coherence of gesture recognition, especially in dynamic environments.

The progress in the development of cutting-edge gesture recognition sensors has found common use in the control of robotic hands. Robotics has increasingly become a substitute for human labor across a broad spectrum of fields[1]. One noteworthy example is the Leap Motion Controller, which generates a virtual screen for visualizing hand gestures and subsequently transmits this information to a robotic hand through wireless device support[2, 3]. Beyond gesture recognition, such technology also finds applications in computer vision tasks[4], robotic control[5, 6, 7, 8], and EEG-based BCI systems[9, 10, 11]. These systems leverage data from MEMS accelerometers, EMG sensors, gyroscopes, flex sensors, and pressure sensors[12, 13], and other sensorial devices [14, 15] enabling the accurate detection and interpretation of human hand movements to facilitate seamless interaction between users and robotic systems. The Leap Motion Controller (LMC) translates human intentions into robotic interactions, excelling at discerning human gestures through hand and finger movements[16, 17, 18]. LMC supports diverse virtual tasks, such as engaging in virtual games, making selections, exercising virtual control, and running applications in virtual environments[19, 20]. Its precision in tracking intricate hand and finger movements within three-dimensional space enhances user engagement and immersion within augmented reality (AR) environments. By enabling lifelike representations of user hands in AR, Leap Motion technology facilitates intuitive interaction with virtual elements. In this paper, we focus on real-time hand tracking using a Leap Motion Controller. We used the Leap Motion Unity Module, a software development kit (SDK) provided by Leap Motion, to integrate hand tracking and gesture recognition into Unity-based applications. We created multiple virtual scenes to enable interactive and immersive real-time experiences. The paper's outline is as follows: Section 2 reviews recent research on hand tracking using Leap Motion Controller and its intersection with computer vision, robotic control, and EEG-based BCI systems. Section 3 describes our proposed methodology for hand tracking across different scenarios. Section 4 outlines the experimental setup and presents results. Section 5 concludes the paper.

2. Related Work

Hand tracking using the Leap Motion Controller has been extensively researched in virtual reality, augmented reality, human-computer interaction, and adjacent fields. For example, Shao et al.[21] leveraged Leap Motion's advanced hand-tracking and gesture-recognition capabilities to enhance interactions within virtual environments. Their study involved configuring hardware, developing interfaces, implementing gesture recognition algorithms,

ICYRIME 2025: 10th International Conference of Yearly Reports on Informatics, Mathematics, and Engineering. Czestochowa, January 14-16, 2025

roumaissa.bekiri@univ-biskra.dz (R. Bekiri); mc.babahenini@univ-biskra.dz (M. C. Babahenini)

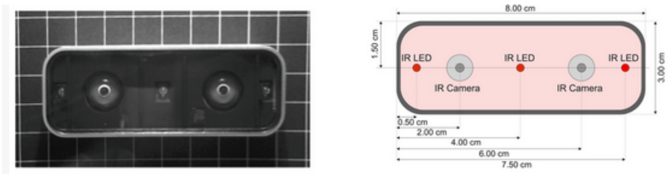

Figure 1: Leap Motion internal structure.

and evaluating system performance through user testing. Satheeshkumar et al.[22] introduced a touchless communication system using Leap Motion, combining virtual environments with real-time settings by interfacing the device with a robotic arm through a microcontroller. Similarly, Sharma et al.[23] demonstrated Leap Motion's ability to track hand movements using infrared LEDs and cameras, offering valuable support for tasks like 3D modeling and animation. In another study, Marin et al.[24] focused on static gesture recognition using Leap Motion and Kinect devices. They designed custom features based on fingertip positions and orientations and utilized a multi-class SVM classifier for gesture classification. Depth features from Kinect were combined with Leap Motion data to improve recognition accuracy. However, their work excluded dynamic gestures. Cooper et al.[25] employed a color-coded glove for hand tracking, which, while effective, reduced user experience due to the need for wearing gloves. In addition to Leap Motion applications, recent advancements in computer vision, such as convolutional neural networks (CNNs)[26, 27, 28], have significantly contributed to improving gesture recognition systems. These techniques are widely adopted in robotic control[29, 30], enabling precise manipulation tasks. Furthermore, EEG-based BCIs[31, 32, 33] have opened new avenues for integrating brain activity signals with gesture control systems, enhancing the interaction between humans and machines. In this paper, we focus on using the Leap Motion Controller for real-time hand tracking and position monitoring, utilizing its SDK to explore applications in dynamic scenarios.

3. System Overview

In our paper, we propose a method using a Leap motion controller to detect and track the human hand using a Software Development Kit(SDK) enabled by Unity 3D. The following are the main steps of our proposed approach:

- A. **Real Hand:** In this initial step, the camera of the Leap Motion Controller(LMC) captures real hand and finger gestures. These captured gestures are then transmitted to a laptop or PC for processing and analysis.
- B. **Virtual Hand:** Subsequently, on the laptop or PC, a 3D image, representing a virtual hand, is generated. This virtual hand is accompanied by relevant information and data associated with its movements and positions. This visual and informational representation of the virtual hand is then wirelessly transmitted to another component or device for further use.
- C. **Interaction with the virtual hand:** In Unity 3D, we create three scenes with virtual objects, the hand can grasp and move in different positions.

3.1. Hardware

The Leap Motion controller, when used in conjunction with the current Application Programmer Interface (API), provides positional data in Cartesian space for predefined objects such as fingertips and pen tips. These positional values are given relative to the central point of the Leap Motion controller, which is situated at the position of the second infrared emitter. As depicted in Figure 2, the controller comprises three infrared emitters and two infrared cameras, classifying it as an optical tracking system based on Stereo Vision. Due to the absence of a point cloud for the scene and the presence of predefined detectable objects, traditional calibration methods are not suitable for Leap Motion. Nevertheless, it is essential to establish a precise reference system for evaluating the accuracy and consistency of the Leap Motion controller. The information is subsequently transmitted to the Leap Motion tracking software via a USB connection. This data materializes as a grayscale representation resembling a stereo image of the nearby infrared light spectrum, captured separately by the left and right cameras.

3.2. Software

To utilize the Leap Motion Controller for hand tracking, we need specialized software. The Leap Motion Controller's Software Development Kit (SDK) is available in various programming languages, including C++, Java, JavaScript, Objective C, C#, and Python. It offers compatibility with different operating systems such as Windows, Linux, and macOS. The benefit of this versatility is that it enables the development of custom software applications capable of tracking human hand movements, empowering developers to create their own unique software solutions tailored to their specific needs.

Following the transmission of photo data to your computer, the real computational heavy lifting takes place. Contrary to common misconceptions, the Leap Motion Controller does not directly generate a depth image; rather, it employs sophisticated algorithms to process the raw sensor data.

3.3. Image captured from LMC

The Leap Motion Controller captures images through its built-in camera arranged in the configuration of an inverted pyramid shape[34]. This controller operates at a frame rate of 120 frames per second, allowing for rapid and continuous image acquisition. Moreover, it offers a wide field of view, spanning 135 degrees, which ensures comprehensive coverage for tracking hand movements and gestures.

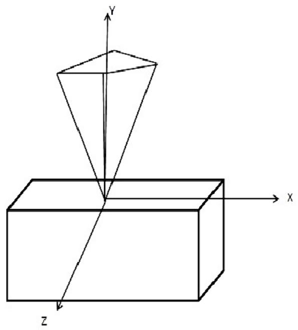

Figure 2: LMC and coordinates.

3.4. Hand skeletal

As mentioned in Figure 3, hand skeletal. It consists of a total of 27 bones. Eight are located in the wrist region, known as the carpals. The palm, referred to as the metacarpals, contains five bones, with each digit being associated with one metacarpal. The remaining fourteen bones are referred to as digital bones, which make up the fingers and thumb.

Within the palm, there are five metacarpal bones, each featuring a head, a shaft, and a base. These metacarpals correspond to the five digits of the hand.

In the fingers and thumb, there are a total of fourteen digital bones, also known as phalanges or phalanx bones. The thumb differs from the other fingers as it lacks a middle phalanx. In contrast, the four fingers possess three phalanges each: the distal phalanx, the middle phalanx, and the proximal phalanx.

Figure 3: Anatomy of Hand Finger Bones.

Additionally, there are small ossified nodes known as sesamoid bones, embedded within the tendons to provide added leverage and alleviate pressure on underlying tissues. These sesamoid bones are typically found at the bases of the digits around the palm, although the precise number can vary among individuals. In this study, the leap motion controller utilizes a conventional skeletal model for tracking the human hand.

3.5. Convert captured human hand into virtual hand

The Leap Motion Controller functions by using human hand motion as input. It achieves this by employing an infrared(IR) sensor integrated into the Leap Motion Controller to detect the motion of the human hand. Additionally, two cameras within the Leap Motion Controller capture the hand movements that are detected by the IR sensors. These hand gestures are then transmitted to the computer through a USB communication port[35]. As depicted in Figure 4. On the computer, there is a Software Development Kit(SDK) that processes the data related to hand gestures. This processing is typically carried out using the Visual C++ programming language, allowing for the manipulation of hand gesture data for virtual simulation. This simulated hand gesture information is then relayed from the computer to a robot hand using a ZigBee wireless device.

4. Experimental Results:

To implement this approach, we construct three distinct virtual environments that are seamlessly integrated with

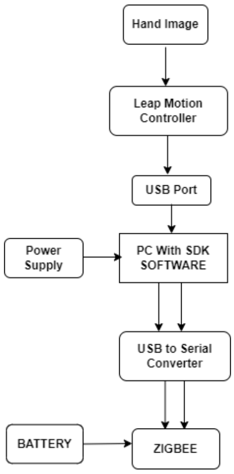

Figure 4: Hand Tracking using Leap motion controller.

the Leap Motion Controller. The integration between the LMC and these virtual environments is crucial for enabling precise and intuitive hand tracking within the digital space. By combining the tracking abilities of the LMC with the interactive capabilities of Unity 3D, we aim to create immersive and interactive experiences that bridge the gap between the physical and digital worlds. We create three virtual environments that interact in real-time with virtual objects. The first scenario is composed of virtual hands that connect with real hands captured from LMC cameras and show the outputs in unity3D cameras. Figure 6. shows that dynamic hand movements rely on fingertip and palm velocity to detect complex movement patterns, contrasting with simpler static gestures. Our approach starts by analyzing global hand movements, including translation, rotation, and circular motion. We focus on individual finger movements, with particular emphasis on the index finger, which holds significant importance in communication and interactions.

5. Conclusion

In this research work, hand tracking and its interaction in different scenarios in real-time. LMC observes hand and finger gestures. It displays a three-dimensional image of a hand. The proposed concept of a linked virtual environment with real-time the environment is found to perform better.

In future, the latency of leap motion controller can be reduced by the new version of SDK. The speed of wireless communication might be improved by the new types of wireless communication. The proposed concept can be implemented in remote-operated robotic applications.

6. Declaration on Generative AI

During the preparation of this work, the authors used ChatGPT, Grammarly in order to: Grammar and spelling check, Paraphrase and reword. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the publication's content.

References

[1] A. Saudabayev, H. A. Varol, Sensors for robotic hands: A survey of state of the art, IEEE Access 3 (2015) 1765–1782.

[2] J. Kofman, X. Wu, T. J. Luu, S. Verma, Teleoperation of a robot manipulator using a vision-based human-robot interface, IEEE transactions on industrial electronics 52 (2005) 1206–1219.

[3] R. Avanzato, C. Randieri, Advances and recent applications of 5g: the model of 5g infrastructure and ecosystem in india, in: CEUR Workshop Proceedings, volume 3869, 2024, p. 78 – 83.

[4] A. Tibermacine, S. M. Amine, An end-to-end trainable capsule network for image-based character recognition and its application to video subtitle recognition., ICTACT Journal on Image & Video Processing 11 (2021).

[5] V. Marcotrigiano, G. D. Stingi, S. Fregnan, P. Magarelli, P. Pasquale, S. Russo, G. B. Orsi, M. T. Montagna, C. Napoli, C. Napoli, An integrated control plan in primary schools: Results of a field investigation on nutritional and hygienic features in the apulia region (southern italy), Nutrients 13 (2021). doi:10.3390/nu13093006.

[6] C. Napoli, V. Ponzi, A. Puglisi, S. Russo, I. Tibermacine, et al., Exploiting robots as healthcare resources for epidemics management and support caregivers, in: CEUR Workshop Proceedings, volume 3686, CEUR-WS, 2024, pp. 1–10.

[7] N. Boutarfaia, S. Russo, A. Tibermacine, I. E. Tibermacine, Deep learning for eeg-based motor imagery classification: Towards enhanced human-machine interaction and assistive robotics 3695 (2023) 68 – 74.

[8] A. Tibermacine, N. Djedi, Gene regulatory network to control and simulate virtual creature's locomotion (2015).

[9] V. Ponzi, S. Russo, V. Bianco, C. Napoli, A. Wajda, Psychoeducative social robots for an healthier lifestyle using artificial intelligence: a case-study, in: CEUR Workshop Proceedings, volume 3118, 2021, p. 26 – 33.

[10] R. Brociek, G. D. Magistris, F. Cardia, F. Coppa, S. Russo, Contagion prevention of covid-19 by

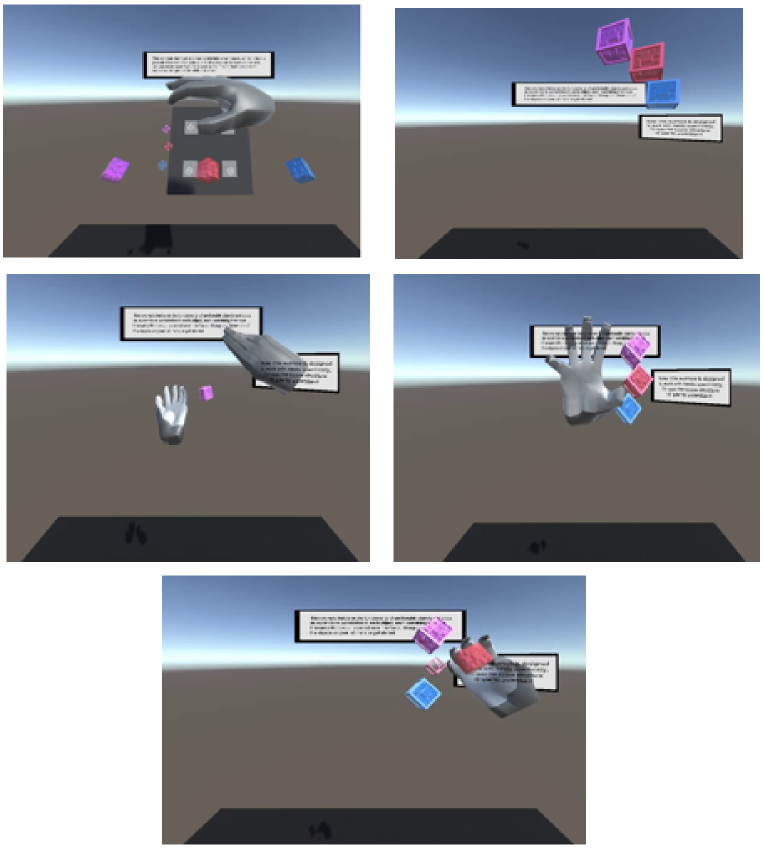

Figure 5: The first scenarios: Top right: Starting tracking. The first line: grasping cubes, Turn right. The second line: Move forward. Moving backward. The last line: interacting with the cube in different situations.

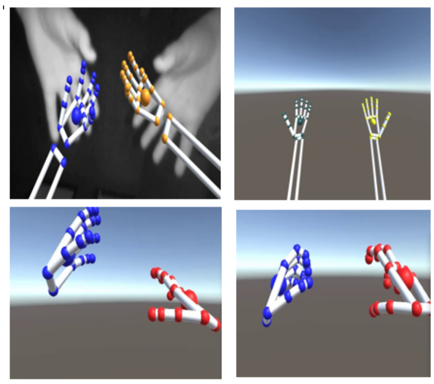

Figure 6: The second scenario: Example of static hand gesture.

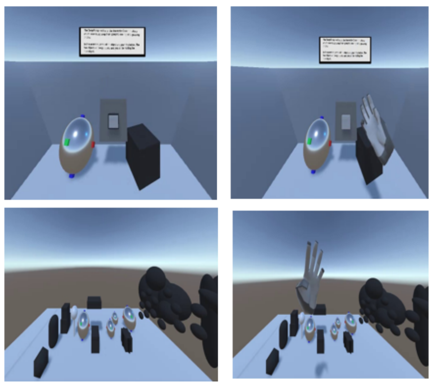

Figure 7: The third scenario: Example of a virtual scene composed of buttons interacting with the virtual hand.

means of touch detection for retail stores, in: CEUR Workshop Proceedings, volume 3092, 2021, p. 89 – 94.

[11] S. Russo, I. E. Tibermacine, A. Tibermacine, D. Chebana, A. Nahili, J. Starczewscki, C. Napoli,

Analyzing eeg patterns in young adults exposed to different acrophobia levels: a vr study, Frontiers in Human Neuroscience 18 (2024) 1348154.
[12] N. S. Chu, C.-L. Tai, Real-time painting with an expressive virtual chinese brush, IEEE Computer Graphics and applications 24 (2004) 76–85.
[13] R. Satheeshkumar, Real time robotic arm control using human hand gesture measurement, Journal of Advanced Research in Dynamical and Control Systems 12 (2020) 984–996.
[14] G. Zimatore, C. Serantoni, M. C. Gallotta, L. Guidetti, G. Maulucci, M. De Spirito, Automatic detection of aerobic threshold through recurrence quantification analysis of heart rate time series, International Journal of Environmental Research and Public Health 20 (2023). doi:10.3390/ijerph20031998.
[15] G. Zimatore, M. Cavagnaro, Recurrence analysis of otoacoustic emissions, Understanding Complex Systems (2015) 253 – 278. doi:10.1007/978-3-319-07155-8_8.
[16] D. Kruse, J. T. Wen, R. J. Radke, A sensor-based dual-arm tele-robotic system, IEEE Transactions on Automation Science and Engineering 12 (2014) 4–18.
[17] C. Randieri, A. Pollina, A. Puglisi, C. Napoli, Smart glove: A cost-effective and intuitive interface for advanced drone control, Drones 9 (2025). doi:10.3390/drones9020109.
[18] A. Tibermacine, I. E. Tibermacine, M. Zouai, A. Rabehi, Eeg classification using contrastive learning and riemannian tangent space representations, in: 2024 International Conference on Telecommunications and Intelligent Systems (ICTIS), IEEE, 2024, pp. 1–7.
[19] H. Jiang, B. S. Duerstock, J. P. Wachs, A machine vision-based gestural interface for people with upper extremity physical impairments, IEEE Transactions on Systems, Man, and Cybernetics: Systems 44 (2013) 630–641.
[20] A. Tibermacine, D. Akrour, R. Khamar, I. E. Tibermacine, A. Rabehi, Comparative analysis of svm and cnn classifiers for eeg signal classification in response to different auditory stimuli, in: 2024 International Conference on Telecommunications and Intelligent Systems (ICTIS), IEEE, 2024, pp. 1–8.
[21] L. Shao, Hand movement and gesture recognition using leap motion controller, Virtual Reality, Course Report (2016).
[22] R. Satheeshkumar, R. Arivoli, Real time virtual human hand using leap motion controller, International Journal of Electronics and Telecommunications 67 (2021).
[23] A. Sharma, A. Yadav, S. Srivastava, R. Gupta, Analysis of movement and gesture recognition using leap motion controller, Procedia computer science 132 (2018) 551–556.
[24] G. Marin, F. Dominio, P. Zanuttigh, Hand gesture recognition with jointly calibrated leap motion and depth sensor, Multimedia Tools and Applications 75 (2016) 14991–15015.
[25] T. B. Moeslund, A. Hilton, V. Krüger, L. Sigal, Visual analysis of humans, Springer, 2011.
[26] B. Ladjal, I. E. Tibermacine, M. Bechouat, M. Sedraoui, C. Napoli, A. Rabehi, D. Lalmi, Hybrid models for direct normal irradiance forecasting: A case study of ghardaia zone (algeria), Natural Hazards 120 (2024) 14703–14725.
[27] E. Iacobelli, D. Pelella, V. Ponzi, S. Russo, C. Napoli, et al., A fast and accessible neural network based eye-tracking system for real-time psychometric and hci applications, in: CEUR WORKSHOP PROCEEDINGS, volume 3870, CEUR-WS, 2024, pp. 32–41.
[28] A. TIBERMACINE, W. GUETTALA, I. E. TIBERMACINE, Efficient one-stage deep learning for text detection in scene images., Electrotehnica, Electronica, Automatica 72 (2024).
[29] A. Tibermacine, N. Djedi, Neat neural networks to control and simulate virtual creature's locomotion, in: 2014 International Conference on Multimedia Computing and Systems (ICMCS), IEEE, 2014, pp. 9–14.
[30] G. Capizzi, C. Napoli, S. Russo, M. Woźniak, Lessening stress and anxiety-related behaviors by means of ai-driven drones for aromatherapy, in: CEUR Workshop Proceedings, volume 2594, 2020, p. 7 – 12.
[31] S. Russo, S. Ahmed, I. E. Tibermacine, C. Napoli, Enhancing eeg signal reconstruction in cross-domain adaptation using cyclegan, in: 2024 International Conference on Telecommunications and Intelligent Systems (ICTIS), IEEE, 2024, pp. 1–8.
[32] I. Naidji, A. Tibermacine, W. Guettala, I. E. Tibermacine, et al., Semi-mind controlled robots based on reinforcement learning for indoor application., in: ICYRIME, 2023, pp. 51–59.
[33] N. Brandizzi, V. Bianco, G. Castro, S. Russo, A. Wajda, Automatic rgb inference based on facial emotion recognition, in: CEUR Workshop Proceedings, volume 3092, 2021, p. 66 – 74.
[34] A. Sarkar, K. A. Patel, R. G. Ram, G. K. Capoor, Gesture control of drone using a motion controller, in: 2016 international conference on industrial informatics and computer systems (ciics), IEEE, 2016, pp. 1–5.
[35] S. Hirche, M. Buss, Human-oriented control for haptic teleoperation, Proceedings of the IEEE 100 (2012) 623–647.