# An Overview of Kinect Based Gesture Recognition Methods

**Alexander Alexeev**
*Institute of Information Technology and Intelligent Systems, Kazan Federal University, Kazan, Russia*

**Tatyana Tsoy**
*Laboratory of Intelligent Robotics Systems (LIRS), Institute of Information Technology and Intelligent Systems, Kazan Federal University, Kazan, Russia*

**Edgar A. Martínez-García**
*Robotics Laboratory, Department of Industrial Engineering and Manufacturing, The Autonomous University of Ciudad Juarez, Chihuahua, Mexico*

**Evgeni Magid**
*Laboratory of Intelligent Robotics Systems (LIRS), Institute of Information Technology and Intelligent Systems, Kazan Federal University, Kazan, Russia*
*HSE University, Moscow, Russia*
*E-mail: san73rus@gmail.com, tt@it.kfu.ru, edmartin@uacj.mx, magid@it.kfu.ru*

**Abstract**

Visual sensors play an important role in a broad variety of robotic systems applications. Even though Kinect technology appeared over 10 years ago, Kinect sensors are still actively employed by researchers around the world. This paper presents an overview of Kinect and Kinect 2 sensors' applications in a human gesture based control. We analyzed existing research papers to estimate a popularity of particular feature extraction and gesture recognition methods, recommendations on a distance between an object of interest and a sensor, reported accuracy and latency of the sensor. Our analysis is supposed to facilitate a selection of a suitable combination of methods for a particular application of Kinect sensor in gesture recognition while considering its performance.

*Keywords*: Microsoft Kinect, robotics, gesture control, gesture recognition.

## 1. Introduction

Robotics has achieved a significant progress, enabling robots to autonomously plan routes [1] and interact with humans in medicine [2], education [3], and rescue operations [4]. Controlling robots with a voice and gestures is the most natural way of interaction [5] due to daily habits of using these methods by humans. Control using gestures can be implemented with vision sensors [6], electromyography methods (that track human muscle contractions [7]), employing a touch screen, an accelerometer or other sensors [8]. This paper overviews gesture control implementations in robotics using Kinect sensor. Kinect is originally a Microsoft motion controller for gaming that uses computer vision methods for a gesture based control. Later Microsoft enabled a custom application development with Software Development Kit for Windows [9]. This study summarize information about Microsoft Kinect use in robotics, compares accuracy of different gesture recognition methods and emphasizes a reported latency of Kinect in various tasks.

## 2. Kinect in Robotics

Kinect sensor (Fig. 1a) is a device introduced by Microsoft in 2010 for XBOX 360 gaming console control. A combination of an infrared sensor and RBG-D camera makes Kinect suitable for human body tracking purposes. Software Development Kit [10] allows developing a custom software for a variety of tasks including gesture based robot control. The introduced by Microsoft in 2013 Kinect v2 was (Fig. 1b) a second generation of Kinect device that also combines RGB-D and infrared sensors. Similar to its predecessor, Kinect v2 is intended for a gesture-based control. However, due to technical improvements such as adding a wide angle time of flight (ToF) camera [11] and upscaling a color camera resolution to 1080p [12], it exhibits a better performance compared to the first generation, which is demonstrated in Section 4.


Fig. 1. Microsoft Kinect first generation and Kinect v2.

Kinect and Kinect v2 capture color and depth data of a scene and are broadly used in computer vision for object recognition including gesture recognition. The distinctive feature of Kinect family is its versatility compared to counterpart sensors. Leap Motion Controller is also used in robotics for gesture-based robot control [13], [14]. Yet, this device is designed to recognize only hand movements, whereas Kinect does not have such limitations and can be employed for a full-body gesture recognition [15] and a face tracking [16]. In addition, Kinect v2 contains an embedded microphone that enables a voice control [17], which can augment a gesture-based control. Leap Motion Controllers could be used in pairs, e.g., to recognize each hand separately [18], which may increase the system's load. Similarly, Kinect sensors can form complicated systems for tracking and motion capture system purposes [19], [20]. However, in some cases Kinect sensor may not be a right solution, e.g., authors in [21] complained that Kinect is not compatible with Linux systems as Kinect's Software Development Kit has only Windows OS support. To solve this issue, the researchers employed Asus Xtion Pro Live with an RGB-D camera similar to Kinect.

In robotics, Kinect could be used to control different types of real robots. In [17] authors presented a Kinect-based industrial robot control and demonstrated that Kinect could be successfully integrated into simulators. In [22] Kinect was applied for a remote control of a mechanical arm that replicates operator-defined gestures. A gesture-based control using Kinect could also be implemented in mobile robotics, where Kinect serves as a robot control panel [23].

## 3. Gesture Recognition Methods

Choosing a gesture recognition method for Kinect based control is important, as it directly affects an overall system performance. If a gesture recognition method is not accurate enough, not suitable for a particular task or a recognition system demonstrates significant delays, a different method of control should be considered. In this section, we summarize information about methods of Kinect based control and emphasize their accuracy.

Table 1 presents an analysis of studies that used various combinations of feature extraction and gesture recognition methods. The first column of the table describes a particular task; the second column *Dist* specifies a recommended optimal distance between an object of interest (gesture producer) and a sensor in meters; the third and the forth – a feature extraction *FEM* and a gesture recognition methods *GRM*, respectively. The fifth column *Acc* shows an accuracy of each approach in percents (as reported by its authors); the sixth *Skeleton* emphasizes if an approach employs a skeleton extraction; the last one *Rf* refers to a corresponding paper. N/A in the table denotes that a paper does not contain corresponding data.

As Table 1 shows, a combination of the Histogram of Oriented Gradients (HOG) for feature extraction and Support Vector Machine (SVM) for gesture recognition ([24], [33]) is a popular approach among researchers. The same combination was used at an earlier stage of [26] research, however, a deep learning AlexNet model eventually was chosen due to a better performance. In [31] the authors emphasized an influence of a distance at which gestures were recognized by Kinect on a recognition accuracy. The highest accuracy of 87% was achieved at a distance of 3 meters from the Kinect, while the lowest accuracy of gesture recognition of 25% was obtained at a distance of 5 meters from the sensor. The importance of a classifier selection when controlling a robot with user-defined gestures was confirmed in [23], where the SURF and FLANN methods were employed, to extract arbitrary points from an image and to recognize gestures. During the experiments, it was discovered that the FLANN library failed to achieve required results in recognizing gestures as it could not match corresponding features in a variety of gestures.

Majority of studies in Table 1 used a method of a skeleton construction, where a skeleton consists of joints of a human body or hands only. This approach allows a more accurate extraction of special points, which are further required for a gesture recognition. An alternative approach built a 3D model of a hand [34]; yet the authors did not provide accuracy indicators for the gesture recognition, and thus a usability of the method for a real gesture recognition system is questionable.

## 4. Kinect Latency

Latency has a significant importance when operating technical devices. Latency can be caused by hardware

Table 1. Gesture Recognition Methods

| Task | Dist, m | FEM | GRM | Acc, % | Skeleton | Rf |
|---|---|---|---|---|---|---|
| Static hand gesture recognition | N/A | Histogram of Oriented Gradients | Support Vector Machine | 98.3 | No | [24] |
| Hand gesture recognition | 1.2–3.5 | Skeleton Extraction | Artificial Neural Network | 97.8 | Yes | [25] |
| Touchless visualization of 3D medical images | 2.5-3.5 | Histogram of Oriented Gradients | AlexNet (CNN) | 96.5 | No | [26] |
| Hand gesture recognition | N/A | MediaPipe Palm Detector | MediaPipe Gesture Recognition | 95.7 | Yes | [27] |
| Sign language recognition for Arabic speakers | N/A | Used (not specified which one) | RandomForest Classifier + Ada-Boosting | 93.7 | Yes | [28] |
| Dynamic hand gesture recognition | N/A | N/A | Dynamic Time Warping | 92 | Kinect embedded | [29] |
| Full body gesture recognition | N/A | CNN | Fast Dynamic Time Warping + CNN | 90.8 | Yes | [30] |
| Smart home control | 3 m | Self-developed | Self-developed | 87 | Yes | [31] |
| Sign language recognition | N/A | Histogram of Oriented Gradients | Dynamic Time Warping | 86 | Yes | [32] |
| Hand gesture recognition | 2.5-3 m | Histogram of Oriented Gradients | Dynamic Time Warping | 76.7 | Yes | [33] |

and software factors. Moreover, a task-dependent software delay sums up with a hardware delay, which leads to a decrease in device performance. This section discusses a latency between a user input and an expected system output, which differ for a particular task.

Table 2. Latency Measurements of Kinect

| Task | Latency, ms | Ref |
|---|---|---|
| **Kinect** | | |
| Skeleton detection | 106-500 | [35] |
| Human fall incident detection | 300 | [36] |
| Contactless hand tracking for surgical robot | 89-576 | [37] |
| Skeleton position estimation | 100 | [38] |
| Skeleton detection | 200-400 | [39] |
| **Kinect v2** | | |
| Human gait analysis | 29-127 | [40] |
| Human gait analysis | 200 | [41] |
| Skeleton detection | 66-100 | [42] |
| IRB4400 industrial robot control | 20 (only a hardware latency) | [43] |
| Rat behavior tracking | 80-126 | [44] |
| 3D content capture | 65-67 | [45] |

Table 2 presents ab information about latency measurements of both generations of Kinect devices employed in various tasks. Note that reference papers of Table 1 and Table 2 are different since Table 1 papers did not specify a latency as they considered a different aspect of Kinect usage.

In [41] authors emphasized a 20 ms hardware latency as a minimum latency level of the device itself, which cannot be reduced. While official manufacture stated latency data for the first generation of Kinect is not available, the study [36] reports the latency value of 100 ms. Based on the surveyed in Table 2 papers we concluded that the average latency for the first generation Kinect is 267 ms, while the average latency for Kinect v2 is 123 ms. However, even though the Kinect v2's average latency significantly improved over the first generation, its effectiveness still depends on a specific task in which it is used and could achieve up to 200 ms.

## 5. Conclusions

This paper presents an overview of Microsoft Kinect sensor applications for gesture based control in robotics. Kinect first generation and Kinect v2 related research papers were analyzed in terms of range and accuracy in a number of typical tasks that require a sensor with RBG-D capabilities. Based on the survey we concluded that the average overall latency of the first generation Kinect is 267 ms, while the average latency for Kinect v2 is 123 ms. However, even though the Kinect v2's average latency has improved over the first generation, its effectiveness naturally depends on a specific task and could achieve up to 200 ms.

**Acknowledgement**

This paper has been supported by the Kazan Federal University Strategic Academic Leadership Program ("PRIORITY-2030").

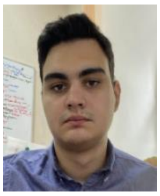
**References**

1. Y. Bai, On Motion Planning and Control for Partially Differentially Flat Systems. Robotica, vol. 39, no. 4, 2021, pp. 718-734.
2. B. Abbyasov, Ultrasound sensor modeling in Gazebo simulator for diagnostics of abdomen pathologies. 2021 International Siberian Conference on Control and Communications (SIBCON), Kazan, 2021, pp. 1-6.
3. L. Gavrilova, Facilitating a preparatory stage of real-world experiments in a humanoid robot assisted English language teaching using Gazebo simulator. 2020 13th International Conference on Developments in eSystems Engineering (DeSE), Liverpool, December 2020, pp. 222-227.
4. E. Magid, Artificial intelligence based framework for robotic search and rescue operations conducted jointly by international teams. Proceedings of 14th International Conference on Electromechanics and Robotics “Zavalishin's Readings” ER (ZR), 2019, pp. 15-26.
5. J. Berger and L. Shuang, Review of Interfaces for Industrial Human-Robot Interaction. Current Robotics Reports, vol. 1, 2020, pp. 27-34.
6. M. Mustafin, Features of Interaction Between a Human and a Gestures-Controlled Collaborative Robot in an Assembly Task: Pilot Experiments. The 2023 International Conference on Artificial Life and Robotics (ICAROB2023), February 2023, pp. 157-161.
7. X. Zhang, A Framework for Hand Gesture Recognition Based on Accelerometer and EMG Sensors. IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans, vol. 41, no. 6, 2011, pp. 1064-1076.
8. S. Berman and H. Stern, Sensors for Gesture Recognition Systems. IEEE Transactions on Systems, Man, and Cybernetics – Part C: Applications and Reviews, vol. 42, no. 3, 2012, pp. 277-290.
9. Kinect – Starting to Develop with Kinect, Available at: https://learn.microsoft.com/en-us/archive/msdn-magazine/2012/june/kinect-starting-to-develop-with-kinect (accessed 29 November 2023).
10. Kinect for Windows, Available at: https://web.archive.org/web/20120103103154/http://www.kinectforwindows.org/ (accessed 10 May 2023).
11. A. Corti, A metrological characterization of the Kinect V2 time-of-flight camera. Robotics and Autonomous Systems, vol. 75, 2016, pp. 584-594.
12. L. Jamhoury, Understanding Kinect V2 Joints and Coordinate System, 2018, Available at: https://lisajamhoury.medium.com/understanding-kinect-v2-joints-and-coordinate-system-4f4b90b9df16 (accessed 29 November 2023).
13. A. Latif, Design and Control Of Autonomous Robot Using Gesture Based Intuitive Interaction. International Journal of Advanced Computer Systems and Software Engineering, vol. 1, no. 4, 2020, pp. 7-11.
14. D. A. Katsay, T. D. Isupova, and A. D. Biryukov, Non-contact control system for anthropomorphic robot. II International scientific and technical conference "Prom-Engineering", May 2016. pp. 441-444.
15. Y. Ma, Deep Learning-Based Upper Limb Functional Assessment Using a Single Kinect v2 Sensor. Sensors, vol. 20, no. 7, 2020,1903 p.
16. L. Vladutu and C. Marian, Framework for Posture and Face Recognition using Kinect an ambient-intelligence method. WSEAS Transactions on Computers, vol. 22, 2023, pp. 13-24.
17. W. Kaczmarek, Industrial Robot Control by Means of Gestures and Voice Commands in Off-Line and On-Line Mode. Sensors, vol. 20, no. 21, 2020, 6358 p.
18. W. Qi, Multi-Sensor Guided Hand Gesture Recognition for a Teleoperated Robot Using a Recurrent Neural Network. IEEE Robotics and Automation Letters, vol. 6, no. 3, 2021, pp. 6039-6045.
19. B. Gabbasov, I. Danilov, I. Afanasyev, and E. Magid, Toward a human-like biped robot gait: Biomechanical analysis of human locomotion recorded by Kinect-based motion capture system. Proceedings of the 10th International Symposium on Mechatronics and its Applications (ISMA15), 2015, pp. 1-6.
20. B. Gabbasov, I. Danilov, I. Afanasyev, and E. Magid, ZMP trajectory from human body locomotion dynamics evaluated by Kinect-based motion capture system. Proceedings of the 11th Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISAPP), 2016, pp. 160-166.
21. N. Pavón-Pulido, IoT Architecture for Smart Control of an Exoskeleton Robot in Rehabilitation by Using a Natural User Interface Based on Gestures. Journal of Medical Systems, vol. 44, 2020, pp. 1-10.
22. P. Tu and C. Huang, Mechanical arm teleoperation control system by dynamic hand gesture recognition based on kinect device. The Journal of Engineering, 2019, vol. 2019, no. 23, 2019, pp. 9110-9113.
23. N. Nikiforov, Pilot studies on Avrora Unior car-like robot control using gestures. Electromechanics and Robotics: Proceedings of 16th International Conference on Electromechanics and Robotics" Zavalishin's Readings" (ER (ZR) 2021), April 2021, pp. 271-283.
24. T. Ma and M. Guo, Research on Kinect-based gesture recognition. 2019 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC), IEEE, September 2019, pp. 1-5.
25. B. Ganguly, Kinect Sensor Based Single Person Hand Gesture Recognition for Man–Machine Interaction. Computational Advancement in Communication Circuits and Systems: Proceedings of ICCACCS 2018, 2020, pp. 139-144.
26. J. Liu, An Improved Kinect-Based Real-Time Gesture Recognition Using Deep Convolutional Neural Networks for Touchless Visualization of Hepatic Anatomical Mode. Journal of Image and Graphics, vol. 7, no. 2, 2019, pp. 45-49.
27. M. Harris, Applying hand gesture recognition for user guide application using MediaPipe. 2nd International Seminar of Science and Applied Technology (ISSAT 2021), Atlantis Press, 2021, pp. 101-108.
28. B. Hisham and A. Hamouda, Supervised learning classifiers for Arabic gestures recognition using Kinect V2. SN Applied Sciences, vol. 1, 2019, pp. 1-21.
29. W. Lv, Gesture recognition in somatosensory game via Kinect sensor. Internet Technology Letters, 2021, 6 p.

30. M. Pfitscher, Activity gesture recognition on kinect sensor using convolutional neural networks and FastDTW for the MSRC-12 dataset. Intelligent Systems Design and Applications: 18th International Conference on Intelligent Systems Design and Applications (ISDA 2018), vol. 1, Springer International Publishing, December 2020, pp. 230-239.
31. H. Fakhrurroja, Hand State Combination as Gesture Recognition using Kinect v2 Sensor for Smart Home Control Systems. 2019 IEEE International Conference on Internet of Things and Intelligence System (IoTaIS), IEEE, November 2019, pp. 74-78.
32. C. Sun, T. Zhang, and C. Xu, Latent support vector machine modeling for sign language recognition with Kinect. ACM Transactions on Intelligent Systems and Technology (TIST), vol. 6, no. 2, 2015, pp. 1-20.
33. K. N. Krisandria, B. S. B. Dewantara, and D. Pramadihanto, HOG-based Hand Gesture Recognition Using Kinect. 2019 International Electronics Symposium (IES), IEEE, September 2019, pp. 254-259.
34. E. E. Shelomtsev, Solution to the problem of gesture recognition for controlling the behavior of a robot. Youth and modern information technologies: collection of papers from the 11th International Scientific and Practical Conference of Students, Postgraduates and Young Scientists, Tomsk Polytechnic University, November 2013, pp. 278-280.
35. M. A. Livingston, Performance measurements for the Microsoft Kinect skeleton. 2012 IEEE Virtual Reality Workshops (VRW), IEEE, March 2012, pp. 119-120.
36. A. A. M. Bigy, Recognition of postures and Freezing of Gait in Parkinson's disease patients using Microsoft Kinect sensor. 2015 7th International IEEE/EMBS Conference on Neural Engineering (NER), IEEE, April 2015, pp. 731-734.
37. Y. Kim, Kinect technology for hand tracking control of surgical robots: technical and surgical skill comparison to current robotic masters. Surgical endoscopy, vol. 28, 2014, pp. 1993-2000.
38. S. Feng, R. Murray-Smith, and A. Ramsay, Position stabilisation and lag reduction with Gaussian processes in sensor fusion system for user performance improvement. International Journal of Machine Learning and Cybernetics, vol. 8, 2017, pp. 1167-1184.
39. S. Park, An approach to the reduction of time delay in a networked performance. Proceedings of the Asia-Pacific Advanced Network, vol. 36, 2013, pp. 91-96.
40. D. Guffanti, The accuracy of the Microsoft Kinect V2 sensor for human gait analysis. A different approach for comparison with the ground truth. Sensors, vol. 20, no. 16, 2020, 4405 p.
41. A. Amini, K. Banitsas, and S. Hosseinzadeh, A New Technique for Foot-Off and Foot Contact Detection in a Gait Cycle Based on the Knee Joint Angle Using Microsoft Kinect v2. 2017 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI), IEEE, February 2017, pp. 153-156.
42. T. Kure and S. Kasahara, Lighterbody: RNN based anticipated virtual body makes you feel lighter. 2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW), IEEE, 2021, pp. 561-562.
43. A. Aalerud and G. Hovland, Evaluation of perception latencies in a human-robot collaborative environment. 2020 IEEE International Conference on Robotics and Automation (ICRA), IEEE, 2020, pp. 5018-5023.
44. J. Taufatofua, P. Pounds, and J. Wiles, PiRat: An autonomous framework for studying social behaviour in rats and robots. 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE, October 2018, pp. 7601-7608.
45. K. Zhou, RGBD-based Real-time Volumetric Reconstruction System: Architecture Design and Implementation. 2022 IEEE International Conference on Visual Communications and Image Processing (VCIP), IEEE, December 2022, pp. 1-5.

## Authors Introduction

Mr. Alexander Alexeev



He received his Bachelor's degree in Regional Studies in 2022 from the Intitute of International Relations, Kazan Federal University (KFU), Russia. He is currently a master student in Institute of Information Technologies and Intelligent Systems (ITIS), KFU, Russia

Professor Edgar A. Martínez-García

He is a full Professor at the Universidad Autónoma de Ciudad Juárez, Mexico; founder and Head of the Robotics Laboratory; leader of the Mechatronics academic body at the Institute of Engineering and Technology, since 2007. He obtained his Ph.D. degree in Robotics Engineering from the University of Tsukuba, Japan (2005).

Ms. Tatyana Tsoy

In 2012 she graduated from International Area Studies Master's Degree Program at the University of Tsukuba. In 2023 completed a PhD in Computer science and information processes at IT IS, KFU.

Professor Evgeni Magid

He is a Head of Intelligent Robotics Department, a founder, and a Head of Laboratory of Intelligent Robotic Systems (LIRS) at Kazan Federal University, Russia. Senior member of IEEE. He obtained Ph.D. degree in Robotics Engineering from the University of Tsukuba, Japan.