iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** I **Month of publication:** January 2024

DOI: https://doi.org/10.22214/ijraset.2024.58038

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue I Jan 2024- Available at www.ijraset.com

A Study on Different Recognition Systems based on Kinect Sensor

Arya J Nair[1], Sandhya S[2], Sreejakumari S[3]
[1]*Dept. of Computer Engineering, NSS Polytechnic College, Pandalam, Kerala, India*
[2]*Dept. of Computer Engineering, NSS Polytechnic College, Pandalam, Kerala, India*
[3]*Dept. of Computer Engineering, NSS Polytechnic College, Pandalam, Kerala, India*

Abstract: ***The Kinect is a motion-sensing gadget that lets us control a computer or gaming console without holding a controller. The influence of Kinect has spread well beyond the video game sector. The Kinect sensor is equipped with multiple innovative sensing modules. It is equipped with four microphones, an RBG camera, and a depth sensor to track body motion, recognize faces, and detect voices. The Kinect sensor's ability to offer depth and RGB-D visual data offers novel avenues in the field of artificial intelligence. The Kinect's adaptable capabilities in-depth perception, skeletal tracking, and voice input have been used for a wide range of interactive experiences, from gaming to healthcare, education, and more. To demonstrate the Kinect's versatility across several domains, the article explores how the technology has been incorporated into gesture recognition, facial recognition, gait, and even activity recognition.***

Keywords: Kinect Sensor, Support vector Machine (SVM) Classifier, Gesture recognition, Face recognition, Human Action Recognition (HAR), Gait recognition, DNM (Discriminant Normal Maps), Convolutional Neural Network (CNN), Variable-length maximal entropy Markov model (MEMM)

I. INTRODUCTION

Microsoft launched the Kinect sensor for depth perception and motion detection. It was initially intended to be used for gaming [1]. Instead of using traditional gaming controllers, Kinect uses motion and voice commands to interact with video games. An RGB color VGA video camera, a depth sensor, and a multi-array microphone are the major components of Kinect. Face and body identification are aided by an RGB camera with 640x480 pixel resolution. The Kinect's depth-sensing technology enables it to identify the location and motion of objects and humans in three dimensions [2]. With the help of speech recognition technology and an array of microphones built to capture ambient sound, users may communicate with the Kinect by giving voice orders to it. Although its initial use was in gaming, the Kinect's technology was also applied in various disciplines, including robotics, object detection, authentication systems, person identification, healthcare, etc. The Kinect sensor detects and records color and depth using infrared light, and then employs advanced algorithms to decipher user movements. In this paper, different recognition systems that are based on Kinect sensors have been presented. An in-depth examination of the mechanisms and functions of numerous systems is provided in Chapter 2, which dives into their complex inner workings. Chapter 4 presents a thorough comparative examination of these various approaches, highlighting their respective advantages, disadvantages, and unique characteristics.

II. REVIEW OF EXISTING APPROACHES

One of the main characteristics that made the Kinect so versatile and useful for interactive applications was its capacity to provide a real-time three-dimensional picture of the surrounding environment. The microphone array of the Kinect makes voice recognition possible [3]. Time-of-flight concepts are used by the depth camera on the Kinect sensor to determine how far an object is from its surroundings. An infrared (IR) projector integrated into the depth camera projects infrared light patterns into the region around it. The depth camera receives reflections of light when it strikes an object. The depth sensor calculates the amount of time it takes for an IR signal to leave the projector, pass through an object, and return to the camera. The time the light signal takes to reflect to the sensor is proportional to the distance between the sensor and the point on the object. Based on the distance information, a depth map is created by Kinect. Skeletal tracking algorithms utilize the depth map to locate and track the position of the joints and body parts [4]. Figure 1 shows the image of the Kinect Sensor.

©IJRASET: All Rights are Reserved



Figure. 1 Kinect Sensor

A. Face Recognition System

A face recognition system based on skeletal tracking of Kinect sensor has been utilized in [5].

The system functions in two phases: the learning phase and the recognition phase. A PC104+ embedded system supports the needs of face detection and recognition techniques. RGB camera in Kinect captures the human image and face detection is performed based on learning-base descriptor and pose-adaptive matching methods using the Microsoft Software Development Kit, which decodes sensor data and recognizes human features in images, a body skeleton may be produced in RBG space. The facial data is extracted from the human image using the head coordinates. Face dimensions in an image are proportional to body distance from the Kinect sensor. An auto-associative neural network has been used for facial recognition tasks. Three elements are required for the neural network to be trained in the cloud: Cloud Blob, Cloud Worker, and Cloud SQL Database. A reliable face identification technique has been shown in [6] using a depth map from Kinect. Its three main sections are canonical preprocessing, discriminant transform, and multi-channel weighted sparse coding. Even if the poses are incorrect, the front view can be inferred from a profile view via canonical preprocessing. Next, RGB is projected into DCS (Discriminant Color Space), and normal maps are projected into DNM (Discriminant Normal Maps) using multi-channel Discriminant Transforms.

B. Gait Recognition System

A pattern of recurring dynamic motions from the body's joints is called a gait. Identifying people through their gait is called gait recognition. In recent years, researchers have been exploring gait recognition to identify people at a distance reliably. A CNN & Kinect sensor-based gait recognition system has been introduced in [7]. Using Kinect, the skeletal sequence is acquired from the human body at the time of walking. Computing the Euclidean distances between two ankles is how gait cycle detection is. The CNN model must extract and train each gait characteristic of every walking sequence. This recognition system has two phases, the registration phase and the identification phase. In the registration phase, the CNN has been used to obtain different gait attributes from skeleton sequences. A 3D matrix made up of joint coordinates during a gait cycle serves as the input for CNN. The CNN is then trained using this input during the registration phase. The identification stage makes use of unknown gait patterns. An individual's identity is predicted using the trained CNN model. Another method utilizes 13 biometric features obtained from Kinect skeleton data for gate recognition in [8]. Three classifiers named 1R, a C4.5 decision tree, and a Naive Bayes classifier have been used in this system. 1R uses just one attribute from the training set to produce a classification algorithm. The statistical independence of the characteristics is taken into account by the Naive Bayes classifier. Several skeletal joints vary in their relative angles (reference point) during human gait. In [9], a gait recognition system based on this angle has been introduced. This study makes use of Kinect V2, which can monitor six individuals at once and 25 skeletal joints per individual.

With this method, individuals may be identified no matter where they are or which way they walk in front of the camera.

C. Human Action Recognition System

Novel concepts in action recognition have been introduced with the release of the Microsoft Kinect depth camera. Kinect-based action recognition involves three steps. It involves skeleton data capturing, action-related feature extraction from skeleton data, and action determination. Kinect uses the depth image to extract skeleton information.

Depth coordinates represent the location of joints in a depth image. After being taken from the skeleton information, the static and dynamic features based on spatial and temporal characteristics are combined. The extracted features are fed to the SVM classifier for action recognition in [10]. The action characteristics are grouped using k-means clustering, and the associations between these clusters are examined using HMMs in [11]. Seven actions were examined in this study. This system can be used for more variant action recognition if more samples train the model. The Variable Length Maximal Entropy Markov Model was used to model a continuous human action recognition system in [12] using RGB depth images acquired from a Kinect sensor. Feature vectors generated based on skeletal information are segmented in real time based on position and motion. The variable-length MEMM technique recognizes human actions according to the results of online model matching.

D. Gesture Recognition Systems

Gesture recognition is a computational approach that identifies and deciphers human gestures. A hand gesture recognition system based on the distance metric Finger-Earth Mover's Distance (FEMD) has been presented in [13]. Initially, the Kinect Sensor is used to gather depth and RBG data. Then hand shape is segmented from the background pixels. The hand gesture recognition process involves dissimilarity measure calculation and template matching. The dissimilarity measure calculation involves finger detection and FEMD calculation. Finger-Earth Mover's Distance (FEMD) has been applied to quantify the differences in hand shapes. Dynamic gesture recognition has been shown in [14].

'Kinect is the data collection tool; it records gesture trajectories in three dimensions. Gesture spotting, data processing, and gesture recognition are the three main processes in the system. The completion of hand gesture is detected by hand close condition. Then orientation features are extracted from the captured data and the SVM classifier is used for gesture recognition.

III. RESULTS AND DISCUSSION

The conventional methods for face identification are extremely complicated and computationally demanding, necessitating robust hardware and capacious, effective memory. A real-time, efficient facial recognition system has been accomplished with the Kinect sensor and Windows Azure cloud in [5]. The Kinect face detection technique outperforms traditional face detection strategies. The system is rotation-invariant, and it does not detect any 2D faces. At a relatively reasonable cost, the system offers high accuracy. Both low-resolution and high-resolution data exhibit equal performance from the DNM-based face recognition method and need only the nose tip's location [6].Kinect's gait-based recognition system achieved higher accuracy in recognition tasks. In [7], the system's accuracy can be further increased by using another CNN architecture with higher performance. Gait recognition based on 1R, Naïve Classifier, and Decision tree is unsuitable for identifying specific individuals in groups since it requires accurate extraction of numerous limb lengths [8]. The joint angle-based recognition system is Scale and perspective invariant [9].Out of the three action recognition methods, an online model matching method that utilizes a Kinect sensor and Variable length Markov Model provides higher accuracy and efficiency [12]. In contrast to alternative techniques for recognizing human actions, the suggested algorithm eliminates the requirement for pre-identifying the beginning and ending points of every human activity.The gesture recognition system based on the FEMD metric obtained high accuracy and high efficiency [13]. It resists localized distortions due to its part-based representation, which allows it to calculate global features. SVM provides higher real-time consistency for gesture detection than HMM, which makes it more useful for a variety of real-world applications.

No	Title /Method	Tool/ Algorithms	Dataset	Accuracy
1	A face recognition system based on a Kinect sensor and Windows Azure cloud technology[5]	Kinect Sensor Windows Azure AANN	Created database	93%
2	Robust RGB-D face recognition using Kinect sensor[6]	Kinect Sensor DCS DNM	CurtinFaces Bosphorus, CASIA FRGC v2 databases	98.4% 97.6% 95.6% 95.2%

Table 1: Face recognition systems based on Kinect

No	Title/Method	Tools/Algorithms	Dataset	Accuracy
1	KinectGaitNet: Kinect-Based Gait Recognition Using Deep Convolutional Neural Network [7]	Kinect V1 Deep CNN, Hierarchical feature extraction	UPCV KGB	96.91% 99.33%
2	Gait Recognition with Kinect [8]	Kinect Sensor 1R, C4.5 decision tree, Naive Bayes classifier	Created dataset	62.7% 76.1% 85.1%
3	Kinect-Based Gait Recognition Using Sequences of the Most Relevant Joint Relative Angles[9]	Kinect v2 DTW-kernel	Kinect skeletal gait database	93.3%

Table 2: Gait recognition systems based on Kinect

No	Title/ Method	Tool/Algorithm	Dataset	Accuracy
1	Human action recognition based on Kinect [10]	Kinect Sensor, SVM classifier	MSR Daily Activity 3D data set	78.8%
2	Human Action Recognition Based on Depth Images from Microsoft Kinect [11]	Kinect sensor, K-means clustering, Hidden Markov Models (HMMs)	Depth images from Kinect sensor	91.4%
3	An Online Continuous Human Action Recognition Algorithm Based on the Kinect Sensor [12]	Kinect Sensor, Variable-length maximal entropy Markov model (MEMM)	Cornell CAD-60 dataset and MSR Daily Activity 3D dataset	92%

Table 3: Action recognition systems based on Kinect

No	Title/ Method	Tool/Algorithm	Dataset	Accuracy
1	Robust Part-Based Hand Gesture Recognition Using Kinect Sensor[13]	Kinect Sensor Finger-Earth Mover's Distance	Created dataset using Kinect Sensor	93.2%
2	A real-time dynamic hand gesture recognition system using kinect sensor[14]	Kinect Sensor v2 SVM	Arabic numbers(0-9) and 26 character dataset	95.42%

Table 4: Gesture recognition systems based on Kinect

IV. CONCLUSION AND FUTURE WORK

With the release of the Kinect, traditional gaming input techniques underwent a dramatic change, opening up new avenues for interactive computer research.

The key features of the Kinect Sensor are motion sensing, skeleton tracking, and voice recognition. Because of its adaptability and capacity to record intricate details about the real world, Kinect technology has been used in many different contexts, fostering innovation and research in the field of interactive technologies.

The study of these identification algorithms reveals the Kinect sensor's wider applications, indicating that its effect extends beyond the gaming industry and into domains including augmented reality, robotics, and healthcare. Kinect's legacy continues to live on in innovative applications and solutions that it sparked.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

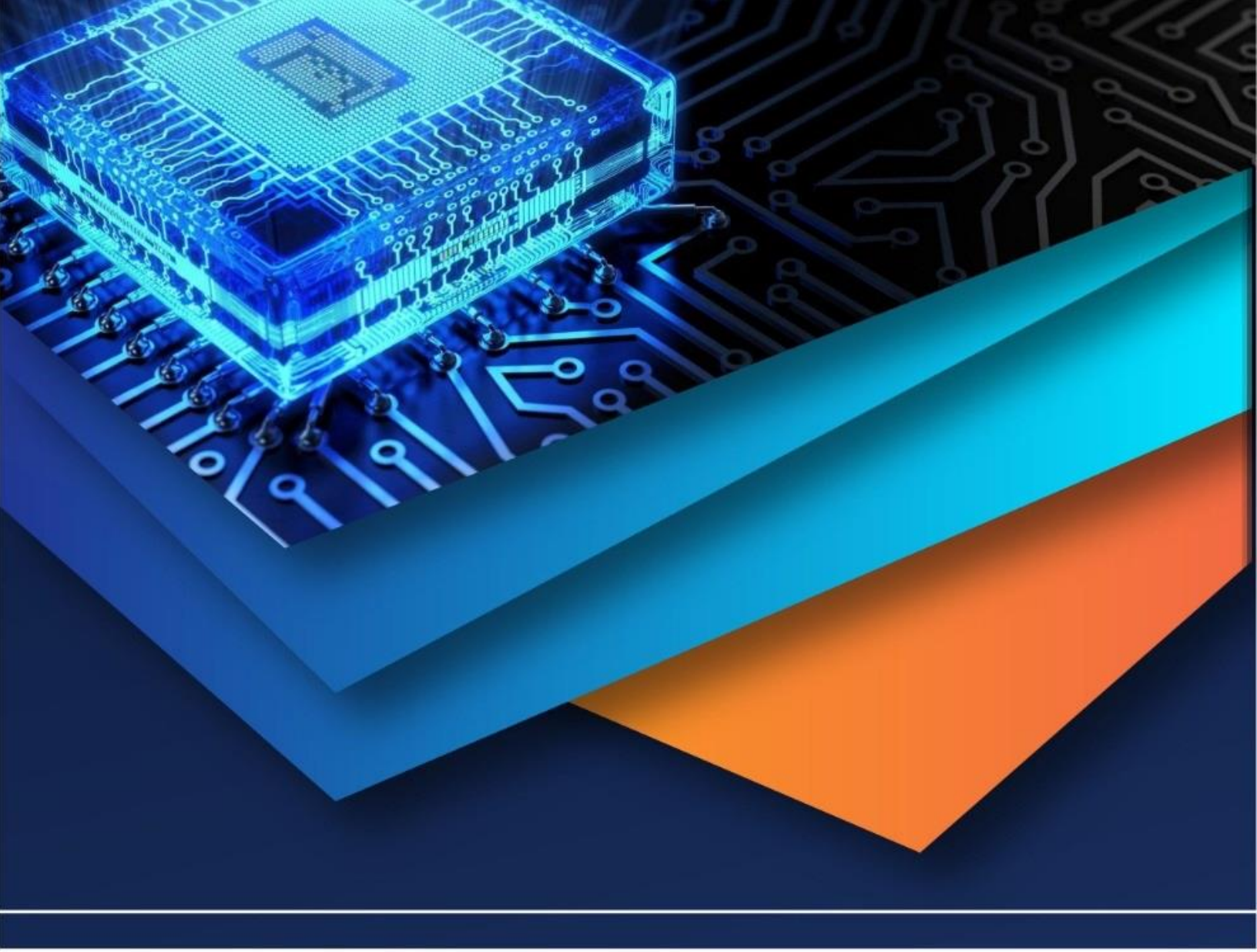
A. Conflict of Interest

The authors have no conflicts of interest to declare.

REFERENCES

[1] https://en.wikipedia.org/wiki/Kinect
[2] https://analyticsindiamag.com/kinect-sensor-the-ai-tool-you-did-not-know-you-had/
[3] https://www.wired.com/2010/11/tonights-release-xbox-kinect-how-does-it-work/
[4] https://learn.microsoft.com/en-us/archive/msdn-magazine/2012/november/kinect-3d-sight-with-kinect
[5] D. -M. Dobrea, D. Maxim and S. Ceparu, "A face recognition system based on a Kinect sensor and Windows Azure cloud technology," International Symposium on Signals, Circuits and Systems ISSCS2013, Iasi, Romania, 2013, pp. 1-4, doi: 10.1109/ISSCS.2013.6651227
[6] Billy Y.L. Li, Mingliang Xue, Ajmal Mian, Wanquan Liu, Aneesh Krishna, Robust RGB-D face recognition using Kinect sensor, Neurocomputing, Volume 214, 2016, Pages 93-108, ISSN 0925-2312, https://doi.org/10.1016/j.neucom.2016.06.012.
[7] Bari ASMH, Gavrilova ML. KinectGaitNet: Kinect-Based Gait Recognition Using Deep Convolutional Neural Network. Sensors. 2022; 22(7):2631. https://doi.org/10.3390/s22072631
[8] Preis, Johannes & Kessel, Moritz & Werner, Martin & Linnhoff-Popien, Claudia. (2012). Gait Recognition with Kinect
[9] Ahmed, Faisal & Paul, Padma Polash & Gavrilova, Marina. (2015). Kinect-Based Gait Recognition Using Sequence of the Most Relevant Joint Relative Angles. Journal of WSCG. 23. 147-156.
[10] Jiahui An et al 2020 J. Phys.: Conf. Ser. **1693** 012190]. In [T. Liu, Y. Song, Y. Gu and A. Li, "Human Action Recognition Based on Depth Images from Microsoft Kinect," 2013 Fourth Global Congress on Intelligent Systems, Hong Kong, China, 2013, pp. 200-204, doi: 10.1109/GCIS.2013.38.
[11] Liu, Tongyang & Song, Yang & Gu, Yu & Li, Ao. (2013). Human Action Recognition Based on Depth Images from Microsoft Kinect. Proceedings - 2013 4th Global Congress on Intelligent Systems, GCIS 2013. 200-204. 10.1109/GCIS.2013.38.
[12] Zhu G, Zhang L, Shen P, Song J. An Online Continuous Human Action Recognition Algorithm Based on the Kinect Sensor. Sensors. 2016; 16(2):161. https://doi.org/10.3390/s16020161
[13] Ren, Z., Yuan, J., Meng, J., & Zhang, Z. (2013). Robust part-based hand gesture recognition using kinect sensor. IEEE transactions on multimedia, 15(5), 1110-1120.
[14] Chen, Yanmei & Luo, Bing & Chen, Yen-Lun & Liang, Guoyuan & Wu, Xinyu. (2015). A real-time dynamic hand gesture recognition system using kinect sensor. 2026-2030. 10.1109/ROBIO.2015.7419071.
[15] Siddharth S Rautaray and Anupam Agrawal, "Vision based hand gesture recognition for human computer interaction: a survey", Artificial Intelligence Review, vol. 43, no. 1, pp. 1-54, 2015.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)