Journal of Robotics and Control (JRC)
Volume 2, Issue 5, September 2021
ISSN: 2715-5072 DOI: 10.18196/jrc.25113

Effective Human Activity Recognition Approach using Machine Learning

Abdul Lateef Haroon P. S.[1], Premachand D. R.[2]
[1, 2]Associate Professor's Department of ECE, BITM, Ballari, Karnataka, India.
Email: [1] abdulbitm@gmail.com, [2] premchandbitm@gmail.com

***Abstract*— The growing development in the sensory implementation has facilitated that the human activity can be used either as a tool for remote control of the device or as a tool for sophisticated human behaviour analysis. With the aid of the skeleton of the human action input image, the proposed system implements a basic but novel process that can only recognize the significant joints. A template for an activity recognition system is provided in which the reliability of the process of recognition and system quality is preserved with a good balance. The research presents a condensed method of extraction of features from spatial and temporal features of event feeds that are further subject to the mechanism of machine learning to improve the performance of recognition. The criticalness of the proposed study is reflected in the outcomes, which when trained using KNN, show higher accuracy performance. The proposed system demonstrated 10-15% of memory usage over 532 MB of digitized real-time event information with 0.5341 seconds of processing time consumption. Therefore on a practical basis, the supportability of the proposed system is higher. The outcomes are the same for both real-time object flexibility captures and static frames as well.**
***Keywords*— KNN, Human Activity Recognition, SVM, RHA.**

I. INTRODUCTION

An RHA aims to recognize, analyze, and predict human activity like moving, walking, standing, sleeping, etc [1]. The RHA uses RGB data to extract the skeletal joints, depth maps and identify human activity [2]. The applications of the RHA have widespread in many areas, such as in video surveillance, human-computer interaction, military applications, etc [3]. The RHA can be performed at different abstraction levels where the actions can be of motion, gesture or any activity. The motion is an activity of a single person that contains different gestures with temporal ordering while the gesture is the elementary movement of a human body [4]-[6].

The activities in general term include multiple numbers of operations (actions or motion) performed by various persons (example: Playing cricket, football, working on different computers by users). The RHA can be performed by using wearable sensors and external devices [7]-[11]. The conventional approaches of the RHA systems are fixed with the predefined point of interest; hence, the activities are dependent on the user interaction of sensors. The common example of the external sensor devices is Intelligent-Homes. These sensors are meant to recognize complex activities such as washing, eating, etc [12]-[15].

This extracted information is purely dependent on the sensor attached to objects or human and interaction among those. But the installing and managing cost of these sensor devices is quite high. To tackle these challenges, various researches have implemented wearable sensor devices for RHA [13]. In this paper wearable sensors, the human activity attributes (GPS and accelerometers), environmental attributes (humidity and temperature sensors), etc are used. This information helps to provide better recognition of human activity [15]-[19].

The past research studies over RHA has focused on the learning as well as recognition of human activities by using traditional cameras captured video sequences. These video sequences can be encoded with the high texture color information, and it helps in image processing. One of the most widespread challenges for the researchers is that taking pictures three-D motion from common cameras. As human things to do are carried out in 3D space, getting access to 3-D records plays a major function in the activity consciousness process. However, the Depth-Map approach is one of the high-quality techniques utilizing for a profitable human endeavour consciousness system compared with usual approaches, a depth-map strategy proven few benefits in the context of activity consciousness process.

II. CONCEPT OF RHA SYSTEM

A. General RHA System

The sensor(s) play an important role in the detection of human activity in traditional RHA [20]–[23]. Figure 1 demonstrates the cycle of recognition of human activity when a movement of the body is received as input. The sensor(s) collect the information obtained from the movement of the human body, and the recognition engine analyses the data and decides the activity type[24]–[29].

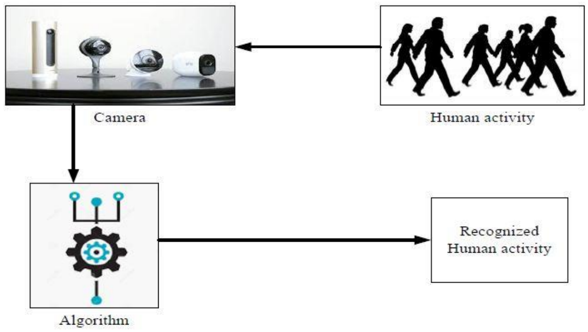


Fig. 1. General Structure of RHA system

To recognize human activities, a depth sensor is used in RHA systems. The depth sensor, in a nutshell, projects infrared beams into the scene and detects them using its infrared detector to determine and measure the depth or distance of each beam from the sensor.

Affective approaches reflect human exercises as per an individual's affective state and passionate communications. Behavioural methods are designed to recognize behavioural attributes, multimodal non-verbal signs for example, motions, outward appearances, and hear-able signs. The phrases "event" and "actions" are typically used in the literature interchangeably. Such words were distinguished in this survey as the sense that the word "event" is used to describe a series of acts that suit specific body movements. On the other hand, the word "actions" is used to describe the activities and events associated with a single person's movements, emotional words, facial expressions, and auditory signs.

III. METHODS

The fundamental methodology used for RHA is to obtain the characteristics from the motion aspect of video or image sequences to facilitate the prediction of one or more specific actions. Nonetheless, one of the larger sets of dependencies associated with human activity identification is the method of extraction of features. Additional capabilities will always improve the recognition of the system's accuracy. Such a mechanism for capturing a higher number of features, however, will require two potential problems, e.g., i) requiring a large amount of processing time and ii) involving additional resources. Both of these two points are harmful to the computational performance of the system. While good progress has been made in research-based techniques related to human activity recognition, it is widely ignored to decide to pick effective features using a cost-effective computational model. Therefore, a system needs to evolve to ensure better quality in recognizing the set of features from the given data. The major research issues are as follows:

- The prior studies do not illustrate the computational difficulties involved with the collection of many features from a given human action data.
- The possibility of minimizing the computational effort of the extraction of features using joint attributes in the skeleton system has attracted less attention among research communities.
- There is currently a very small number of literatures dedicated to the exploration of an effective number of joints responsible for human activity recognition.
- Optimizing the use of depth maps has less research, and more research is aimed at applying machine learning to provide greater accuracy.

The framework takes an input of image sequences with defined standard human actions with the help of analytical research methodology. This process is followed by extracting the depth map from the data set input image, which is resumed to extract the predicted segments on three separate x-y-z axis planes. An algorithm is built that uses different types of identity-based attributes to test all the joints of the depth object skeleton. This method ultimately results in an appropriate number of joints responsible for the actual detection of the successful capture of the mechanism of awareness of human activity. The major contribution of the proposed model is that it delivers faster joint processing regardless of any selection of movement patterns and does not offer any form of dependency on the computation of an unnecessary number of joints to recognize effective human action.

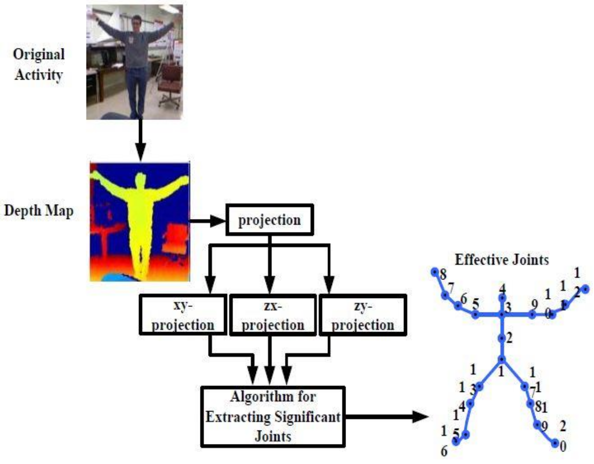


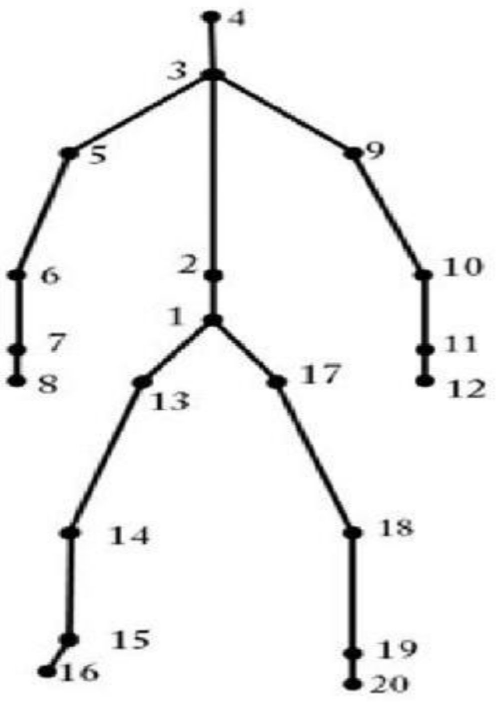
Fig. 2. Proposed Framework

IV. SYSTEM DESIGN

Depth cameras have received significant consideration from researchers in the vision and robotics community because of the financially cost-effective "Kinect." The camera depth has two main advantages. Firstly, to recover postures and recognize the activity, the depth sensor provides information about the 3D image structure. Secondly, the depth sensor can sense in darkness. These benefits are used from a depth map in interesting research points such as human skeleton detection. The skeletons measured from maps of depth are accurate and bring benefits to various applications, including recognition of actions and gestures. Depth map RHA can also be described as a sequence of object representation, extraction of features, and identification of these activities in its simplest form.

The proposed RHA takes advantage of the 3-D skeleton information provided from the 2.0 SDK Kinect. The Kinect 2.0 sensor has a high depth-fidelity that enables it to more clearly observe the tiny objects, resulting in highly accurate 3-D object building. The device can control each person's number of people and 20 skeletal joints. Figure.4.3 shows the skeletal 3-D view joints monitored by the 2.0 sensor Kinect. Many implementations move include the proposed RHA. The first step is to record and track the entire event process using a kinect2.0 sensor through the video sequence. Because activity is the cyclic locomotion, it detects the entire cycle of activity that provides consistent extraction of features. Figure 3 represents the practical implementation process of proposed RHA.

Relative joint angle (RJA) characteristics of various joints are measured over the entire operation period during this process. The major advantage of using RJA is that the process allows invariants to be scaled-up and perceived. Hence, motion recognition is not restricted by the constant distance from the Kinect or single person moving from the Kinect camera to the specific direction. To assess the relevance of the particular activity related RJA function, the device undergoes statistical analysis, which calculates the relative joint pair angles based on activity movement. Here, systems find only the most relative joint angles based on RJA measurement and obtain the characteristics of operation. Once the feature extraction has been completed, the system continues to the classification process (i.e., FITCECOS), which is SVM's "Match multiclass classification models." The proposed RHA considers the collection of most appropriate RJA sequences from both training and testing data samples as parameters and assesses the non-similarity characteristics between them. The major advantage of the proposed RHA is that, due to object walking speed, it can validate the variable size RJA sequences initiated in different videos of the same person (5 data samples were considered in this). Finally, by evaluating skeletal joints, the system recognizes the individual.



1 Hip center	11 Wrist right
2 Spine	12 Hand right
3 Shoulder center	13 Hip left
4 Head	14 Knee left
5 Shoulder left	15 Ankle left
6 Elbow left	16 Foot left
7 Wrist left	17 Hip right
8 Hand left	18 Knee right
9 Shoulder right	19 Ankle right
10 Elbow right	20 Foot right

Fig. 3. Skeletal joints tracked by Kinect sensor

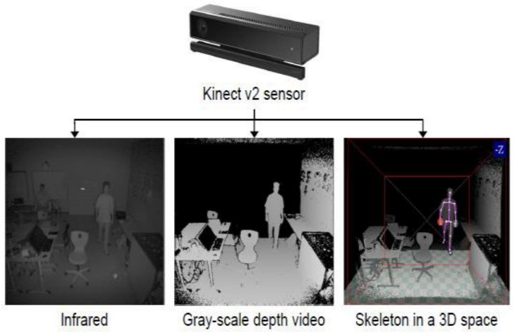


Fig. 4. Proposed model obtains various skeletal data streams

V. RESULTS

The proposed system is applied over various different video sequences samples to perform recognition of variable human activity. In this section sample video sequence is considered to assess the performance and efficiency of the proposed system. The programming is carried out for all human actions including high arm wave, sidekick, hammer, etc. The results of these actions are shown below. The depth silhouettes of the various activity offer the activity's shape details. The human activity of waving high arm is identified in the following Figure 5 in terms of depth silhouettes, and significant joints are tracked or marked. The joints are represented by a number (1, 2, 3, 20). Similarly, the joints in the 3D plane are given for xy, yz, and zx projections

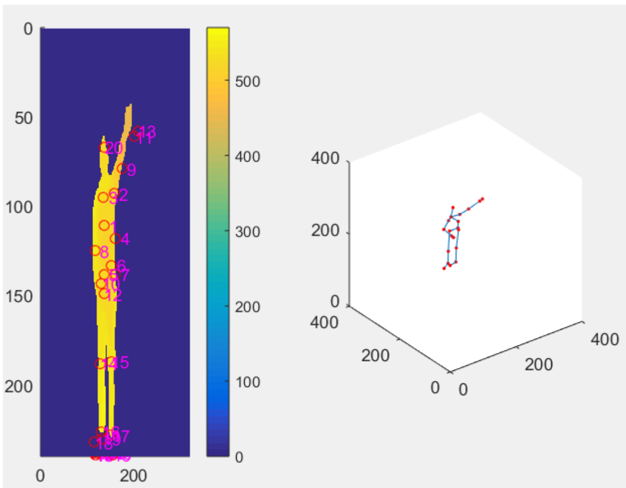


Fig. 5. Depth and Projection analysis of the high arm wave activity

The displacement analysis of the high arm wave activity against different frame is given in Figure 6. The standard deviation of the high arm wave is represented in Figure 7 where the higher deviation of the displacement is found about 10.

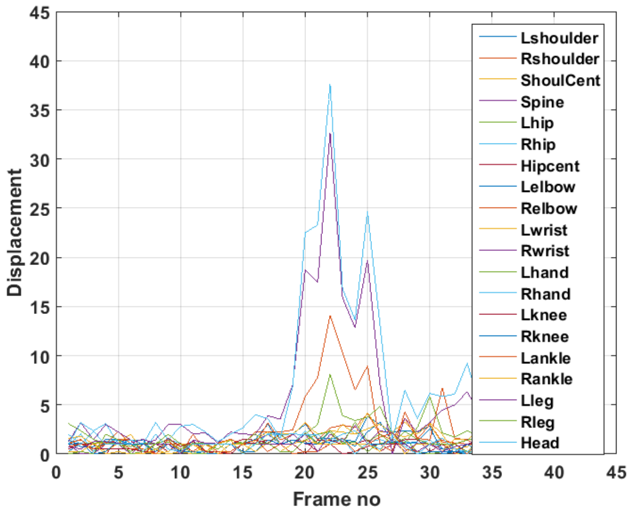


Fig. 6. Displacement at different frame numbers of high arm wave activity.

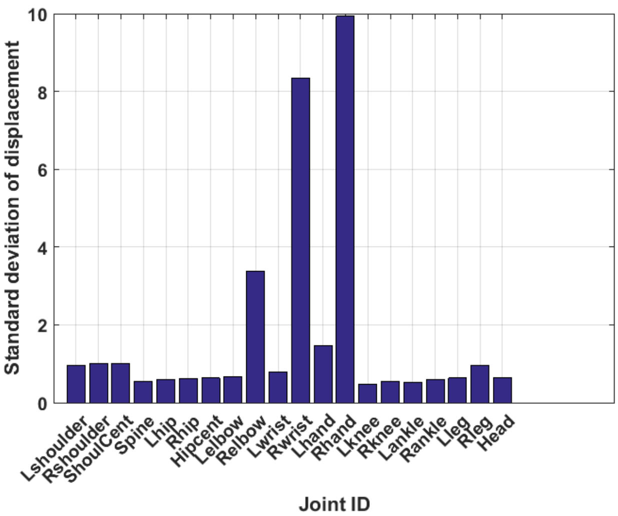


Fig. 7. Standard deviation of displacement at high arm wave activity

VI. CONCLUSION

The combined features of geometry and visual attributes may provide a synchronized platform to provide technology for the human to machine interaction (HMI). The active involvement of the human developed initially many HMI based application into the system, but in the future, much future vision of the smart, intelligent and ubiquitous applications requires the certain context of the humans such as their critical activities as an input. The process of finding a fast, accurate feature set of coordinates is a very challenging task for the accuracy of recognition of the activities. The Recognition of Human Activity (RHA) is an open research issue as there exists an evolving process in both imaging techniques and camera technology. The problem of RHA is portrayed as a "Three-level categorization problem," namely 1) Action primitives, 2) Actions/activities and 3) Interactions. The research study involves the understanding of the various approaches of image representation and classifications for vision-based activity recognition or human behaviour.

REFERENCES

[1] N. Käse, M. Babaee and G. Rigoll, "Multi-view human activity recognition using motion frequency," *2017 IEEE International Conference on Image Processing (ICIP)*, Beijing, 2017, pp. 3963-3967.

[2] D. Tao, L. Jin, Y. Yuan and Y. Xue, "Ensemble Manifold Rank Preserving for Acceleration- Based Human Activity Recognition," in *IEEE Transactions on Neural Networks and Learning Systems*, vol. 27, no. 6, pp. 1392-1404, June 2016.

[3] H. Zhang and L. E. Parker, "CoDe4D: Color-Depth Local Spatio-Temporal Features for Human Activity Recognition From RGB-D Videos," in *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 26, no. 3, pp. 541-555, March 2016.

[4] W. Lu, F. Fan, J. Chu, P. Jing and S. Yuting, "Wearable Computing for the Internet of Things: A Discriminant Approach for Human Activity Recognition," in *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 2749-2759, April 2019.

[5] J. Qi, Z. Wang, X. Lin and C. Li, "Learning Complex Spatio-Temporal Configurations of Body Joints for Online Activity Recognition," in *IEEE Transactions on Human-Machine Systems*, vol. 48, no. 6, pp. 637-647, Dec. 2018.

[6] Abdul Lateef Haroon P.S., U.Eranna, "Insights of Research-based approaches in Human Activity Recognition" Communications on Applied Electronics (CAE), Volume 7- No.16, May- 2018

[7] Abdul Lateef Haroon P.S., U.Eranna, "An Efficient Activity Detection System based on Skeletal Joints Identification" International Journal of Electrical and Computer Engineering (IJECE), Vol.8, No.6, December-2018, pp. 4995~5003.

[8] Abdul Lateef Haroon P.S., U.Eranna, "A Simplified Machine Learning Approach for Recognizing Human Activity" International Journal of Electrical and Computer Engineering (IJECE), Vol.9, No.5, October-2019, pp. 3465~3473.

[9] Abdul Lateef Haroon P.S., U.Eranna, "Modelling Approach to select Potential Joints for Enhanced Action Identification of Human" 3rd IEEE International Conference on Electrical, Electronics, Communication, Computer Technologies and Optimization Techniques (ICEECCOT-2018), 2018.

[10] Abdul Lateef Haroon P.S., U.Eranna, "Human Activity Identification using Novel Feature Extraction and Ensemble-based learning for Accuracy" Artificial Intelligence Methods in Intelligent Algorithms, Proceedings of 8th Computer Science On-line Conference, Vol.2, 2019.

[11] Y. Guo *et al*., "Tensor Manifold Discriminant Projections for Acceleration-Based Human Activity Recognition," in *IEEE Transactions on Multimedia*, vol. 18, no. 10, pp. 1977-1987, Oct. 2016.

[12] L. Peng, L. Chen, X. Wu, H. Guo and G. Chen, "Hierarchical Complex Activity Representation and Recognition Using Topic Model and Classifier Level Fusion," in *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 6, pp. 1369-1379, June 2017.

[13] J. Cheng, H. Liu, F. Wang, H. Li and C. Zhu, "Silhouette Analysis for Human Action Recognition Based on Supervised Temporal t-SNE and Incremental Learning," in *IEEE Transactions on Image Processing*, vol. 24, no. 10, pp. 3203-3217, Oct. 2015.

[14] X. Chang, W. Zheng and J. Zhang, "Learning Person–Person Interaction in Collective Activity Recognition," in *IEEE Transactions on Image Processing*, vol. 24, no. 6, pp. 1905-1918, June 2015.

[15] M. Ehatisham-Ul-Haq *et al*., "Robust Human Activity Recognition Using Multimodal Feature-Level Fusion," in *IEEE Access*, vol. 7, pp. 60736-60751, 2019.

[16] R. Xi *et al*., "Deep Dilation on Multimodality Time Series for Human Activity Recognition," in *IEEE Access*, vol. 6, pp. 53381-53396, 2018.

[17] Z. Yang, O. I. Raymond, C. Zhang, Y. Wan and J. Long, "DFTerNet: Towards 2-bit Dynamic Fusion Networks for Accurate Human Activity Recognition," in *IEEE Access*, vol. 6, pp. 56750- 56764, 2018.

[18] H. Zou, Y. Zhou, R. Arghandeh and C. J. Spanos, "Multiple Kernel Semi-Representation Learning With Its Application to Device-Free Human Activity Recognition," in *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 7670-7680, Oct.

[19] N. Alshurafa *et al*., "Designing a Robust Activity Recognition Framework for Health and Exergaming Using Wearable Sensors," in the *IEEE Journal of Biomedical and Health Informatics*, vol. 18, no. 5, pp. 1636-1646, Sept. 2014.

[20] G. Azkune and A. Almeida, "A Scalable Hybrid Activity Recognition Approach for Intelligent Environments," IEEE Access, vol. 6, pp. 41745–41759, 2018.

[21] C. Wang, Y. Xu, H. Liang, W. Huang, and L. Zhang, "WOODY: A Post-Process Method for Smartphone-Based Activity Recognition," IEEE Access, vol. 6, pp. 49611–49625, 2018.
[22] R. Xi et al., "Deep Dilation on Multimodality Time Series for Human Activity Recognition," IEEE Access, vol. 6, pp. 53381–53396, 2018.
[23] H. Wang, Y. Xue, X. Zhen, and X. Tu, "Domain Specific Learning for Sentiment Classification and Activity Recognition," IEEE Access, vol. 6, pp. 53611–53619, 2018.
[24] A. Yousefzadeh, G. Orchard, T. Serrano-Gotarredona, and B. Linares-Barranco, "Active Perception With Dynamic Vision Sensors. Minimum Saccades With Optimum Recognition," IEEE Trans. Biomed. Circuits Syst., vol. 12, no. 4, pp. 927–939, Aug. 2018.
[25] Y. Hsu, S. Yang, H. Chang, and H. Lai, "Human Daily and Sport Activity Recognition Using a Wearable Inertial Sensor Network," IEEE Access, vol. 6, pp. 31715–31728, 2018.
[26] L. Cai, X. Liu, H. Ding, and F. Chen, "Human Action Recognition Using Improved Sparse Gaussian Process Latent Variable Model and Hidden Conditional Random Filed," IEEE Access, vol. 6, pp. 20047–20057, 2018.
[27] W. Li, B. Tan, and R. Piechocki, "Passive Radar for Opportunistic Monitoring in E-Health Applications," IEEE J. Transl. Eng. Heal. Med., vol. 6, no. December 2017, pp. 1–10, 2018.
[28] Z. Yang, O. I. Raymond, C. Zhang, Y. Wan, and J. Long, "DFTerNet: Towards 2-bit Dynamic Fusion Networks for Accurate Human Activity Recognition," IEEE Access, vol. 6, pp. 56750–56764, 2018.
[29] S. Savvaki, G. Tsagkatakis, A. Panousopoulou, and P. Tsakalides, "Matrix and Tensor Completion on a Human Activity Recognition Framework," IEEE J. Biomed. Heal. Informatics, vol. 21, no. 6, pp. 1554–1561, 2017.