

Detecting Complex 3D Human Motions with Body Model Low-Rank Representation for Real-Time Smart Activity Monitoring System

Ahmad Jalal¹, Shaharyar Kamal² and Dong-Seong Kim³

¹Department of Computer Science & Engineering, Air University
Service Road E-9/E-8, Islamabad 44000 - Pakistan
[e-mail: ahmadjalal@mail.au.edu.pk]

²School of Computer Engineering, Mid-Sweden University
Sundsvall, Holmgatan 10, 85170 - Sweden
[e-mail: sherrykhan84@gmail.com]

³School of Electronic Engineering, Kumoh National Institute of Technology
Yanghოდong, Daehakro 61, 730-701 - South Korea
[e-mail: dskim@kumoh.ac.kr]

*Corresponding author: Ahmad Jalal

*Received June 18, 2017; revised September 18, 2017; accepted October 17, 2017;
published March 31, 2018*

Abstract

Detecting and capturing 3D human structures from the intensity-based image sequences is an inherently arguable problem, which attracted attention of several researchers especially in real-time activity recognition (Real-AR). These Real-AR systems have been significantly enhanced by using depth intensity sensors that gives maximum information, in spite of the fact that conventional Real-AR systems are using RGB video sensors. This study proposed a depth-based routine-logging Real-AR system to identify the daily human activity routines and to make these surroundings an intelligent living space. Our real-time routine-logging Real-AR system is categorized into two categories. The data collection with the use of a depth camera, feature extraction based on joint information and training/recognition of each activity. In-addition, the recognition mechanism locates, and pinpoints the learned activities and induces routine-logs. The evaluation applied on the depth datasets (self-annotated and MSRAction3D datasets) demonstrated that proposed system can achieve better recognition rates and robust as compare to state-of-the-art methods. Our Real-AR should be feasibly accessible and permanently used in behavior monitoring applications, humanoid-robot systems and e-medical therapy systems.

Keywords: 3D human pose estimation, skeleton model, real time system, smart home

This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2017-2014-0-00639) supervised by the IITP (Institute for Information & Communications Technology Promotion).

1. Introduction

Most of the signal/video or image data used nowadays are retrieved by a body-attached/hanging sensors such as body markers or video cameras, which are the most widely used sensors. These sensors recovering the valuable data and advent the automated home monitoring systems in the present digital technologies and intelligent systems era, made the significant contributions to identify the daily routine activities of the inhabitants, create activity log information for health observation and provide support to the problems being faced by the individuals residing independently [1-5]. A routine-log is basically an auto-generated sequential records in which human's daily routine behaviour is monitored, displayed and saved in the form of video or by wearable sensor's signals data [6-8]. It can be implanted/fixed in the user gadgets in numerous applications like elderly healthcare services, security surveillance and smart home [9-12].

To apply routine-log features over activity recognition (AR), there are two key factors to recognize the human activities such as extraction of appropriate features and classification of these extracted features to select the appropriate activity [13]. Therefore, to explore feature extraction method, many previous approaches have been proposed such as silhouette based AR, skeleton-joints based AR and spatiotemporal interest point based AR [14-16]. Whereas, for the classification approach, several linear and probabilistic classifiers have been used in the field of computer vision and pattern recognition to classify the human activities [16-18].

Similarly, we used multiple sensors to measure data for AR in the form of signal or images. In case of RGB (i.e., digital) cameras, features values and classification are just restricted with limited information. In addition, these cameras need complex installation which make high computational cost and unsuitable for the real-time applications. One of the major drawbacks of RGB sensor is that it is hard to differentiate the similar human activities performed by different subjects having similar height and body shape [19]. Therefore, these sensors provide low recognition accuracy and fail to recognize different activities efficiently. Due to above reasoning, we used newly launched depth cameras to recognize human activities. In case of depth camera, it provides maximum pixel information (both intensity plus digital data), less expensive, easy installation, joints body parts information and distance based subject information.

Using multi-camera (i.e., depth) scenarios, we can avoid several issues such as 3D non-rigid reconstruction problem, dynamic background complexities or lighting issues during continuous movements of subject and gain numerous features such as body-postural construction and recognizing the human activities [20-22]. In [23], Xia et al. showed the human posture using histogram of 3D joints locations (HOJ3D). These features are then extracted by transforming the invariance skeleton joints into spherical coordinates to recognize human behavior. In [24], Oreifej and Liu used skeleton joint information to model body pose and motion information. Then, motion and geometry cues based on histogram of normal orientation (i.e., 4D depth, time and spatial coordinates) are successfully processed to recognize the human activities. In [25], Kamal et al. represented the human skeleton by considering joint information (i.e. relative joint positions, temporal movement of joints and offset of the joints) with respect to the depth intensity values to evaluate recognition performance. Until now, to the best of our knowledge, there are very limited work has been exposed in which joints information, depth silhouettes and routine-logging features are used in a single platform. Therefore, we planned to implement a routine-logging system having novel

feature extraction techniques, handling self-occlusion issues, managing human's scaling normalization and utilizing depth data to recognize the daily human activities in real time. In [26], Ming et al. proposed hybrid texture-edge local pattern coding and integration of RGB and depth feature values to examine the recognition rate. In [27], Charalampous and Gasteratos developed on-line deep learning algorithm treated with spatio-temporal features for classification accuracy of different actions.

In this paper, we focus on tackling the problems of tracking human skeleton generation models, complex 3D motions especially for the 3D human postures and treatment with bulk of unnecessary data during continuous routine-log's activity recording. In addition, some useful features are concatenated by two different features types at the extraction level such as position features deal with local motion and directional features that measure the movements of joint points information, respectively. In order to examine the recognition performance, a new depth activity dataset is provided that contains segmented video sequences for training/testing, and ensures that the proposed method can be applied more broadly. Also, it will become a benchmark dataset for routine-log activity recognition based on stereo camera.

In summary, the major contributions of this research paper are shortened as follows: (1) We proposed the real-time AR system combined with likelihood recognizer values that can track the occluded and complex postures. (2) Also, it can detect very fast body parts movements by considering posture normalization mechanism. (3) It has major novelty in the form invariant characteristics along with translation and scaling techniques. Also, training and testing processes consume less computational time and complexity as compare to deep learning [26] and other methods [23, 24]. Some real-time smart activity monitoring systems examples are mentioned in Fig. 1.



Fig. 1. Few examples of real-time smart activity monitoring system environment.

The remainder portion of the paper is arranged as follows. In Section 2, we provide a detailed procedure of the proposed method based on depth data acquisition, feature extraction techniques and recognizer engine. In Section 3, we discuss experimental settings and results. Finally, Section 4 presents the conclusion.

2. Method and Design

2.1 Proposed Routine Logging System

Here, we give a brief introduction of few research items which are most relevant to the proposed method. The data acquisition process (i.e., removing noise, generate stable skeleton models, joint identifications), second is motion feature generation (i.e., position features), third is recognizer engine as HMM for training and real-time testing (i.e., evaluating each activity status) and lastly, routine-logging files are developed to make use of the well-managed life routines of the user. An input sensor is used for the proposed system which is the Microsoft Kinect depth sensor that produces RGB and depth maps based on distance information.

Fig. 1 shows the overall flow of proposed system. The process of training each activity in proposed system is described in **Fig. 1(a)**, while to find maximum likelihood among all activities and to create routine logs from the identified activities is demonstrated in **Fig. 1(b)**.

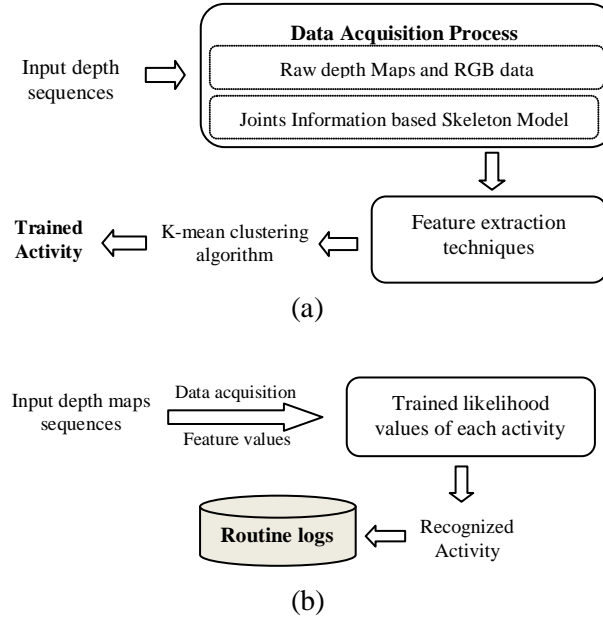


Fig. 1. System flow of the proposed approach. (a) processes of training routine logging system and (b) recognizing activities using the trained HMMs.

2.1.1 Data Acquisition Process

Initially, a succession of depth silhouettes based on routine-log Real-AR system are extracted having user's everyday logs activities via depth sensor. The depth silhouette of the human body is drawn out from each depth data having a size of 320x240. Few samples of depth silhouettes-based taking medicine and hand clapping activities are presented in **Fig. 2**.



Fig. 2. Examples of depth activities silhouettes used in our Real-AR system.

2.1.2 Joints Information based on Skeleton Body Model

A skeleton body model is generated [28] from its corresponding human body silhouette for each activity [29] to express the joints information. **Fig. 3** represents the sample joints based poses with their respective depth silhouettes.

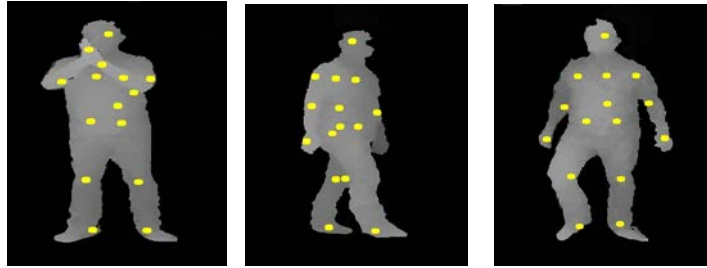


Fig. 3. Various examples of joints position information derived from their corresponding depth silhouette.

The skeleton pose is merged together with the information of joint points [30]. Consequently, the skeleton pose/model is described with the help of the joint points of different body parts. In **Fig. 4(a)**, depth silhouette is presented. While, in **Fig. 4(b)**, a skeleton model is sketched accompanied by a vector of 15 joints of body parts. Whereas, the relative skeleton body model of (a) is depicted in **Fig. 4 (c)**. Circles (i.e., red in color) are used to demonstrate 15 body joints.

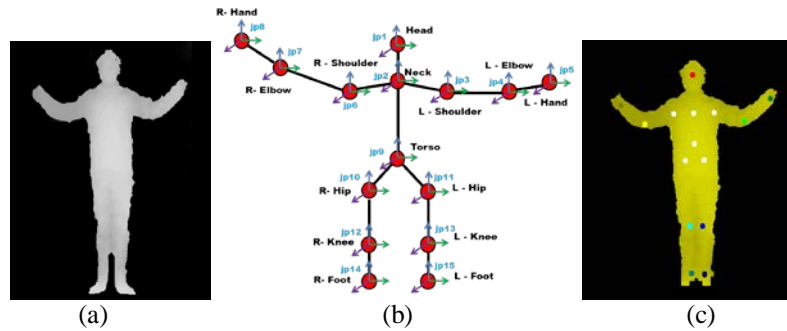


Fig. 4. (a) Depth silhouette, (b) skeleton body part labeling, and (c) skeleton model containing 15 joints.

2.2 Feature Generation Techniques

We execute position motion and difference features [31], after obtaining fifteen 3D coordinates joints values (J_x, J_y, J_z). The position motion M_J of the motion parameters [32] measures the distance between reference and current joints of successive frames [33] and is defined as

$$M_J = \sqrt{(J_{x(i-1)} - J_{x(i)})^2 + (J_{y(i-1)} - J_{y(i)})^2 + (J_{z(i-1)} - J_{z(i)})^2} \quad (1)$$

where J_x, J_y, J_z are the 3D joint structure matrix which denote the 3D position of the i^{th} body joints. The left-sided term in Eq. (1) is the position motion joint coordinate values (x, y, z) of pre-reference frame $J_{(i-1)}$ which encode 3D positions of tracked features points. The right-sided term encourages the 3D joints structures coordinates of reference frame $J_{(i)}$. Furthermore, a set of position motion characteristics are acquired using the sampled activities accompanied by a series of frames (F). In **Fig. 5**, we have three different parameters in which “Magnitude (Position)” calculates the position values of 3D coordinates, “F” represents the total frames used as 0 to 10 and “no. of features” are extracted from 15 body joints areas.

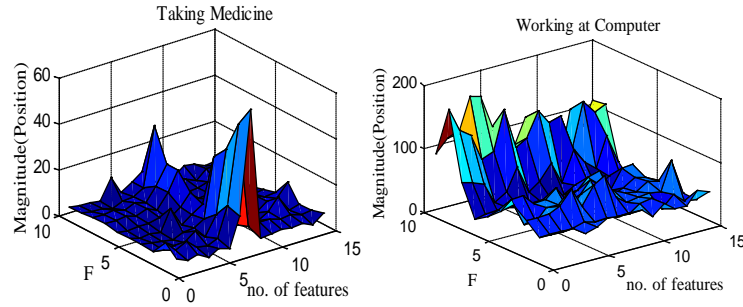


Fig. 5. Position motion features of 15 joint points having different activities used in routine-log Real-AR system.

We measure the movement angles difference of i th body joints between two consecutive frames to formulate the position angles features, which is computed as;

$$D_{J(C_1, C_2)} = \tan^{-1} \left(\frac{JC_{2(i-1)} - JC_{2(i)}}{JC_{1(i-1)} - JC_{1(i)}} \right), \quad (2)$$

where C_1 and C_2 denote the x position of the i th body joints at successive frames respectively and \tan is a vector of 45 dimensions (15×3). Few examples of position angles features by considering *taking medicine* and *working at computer* activities are represented in **Fig. 6**. In **Fig. 6**, “Direction (rad)” measures the angle change in each joint point, “F” represents the total frames and “no. of features” consists of x , y , and z coordinates values at 3D joint points as 45.

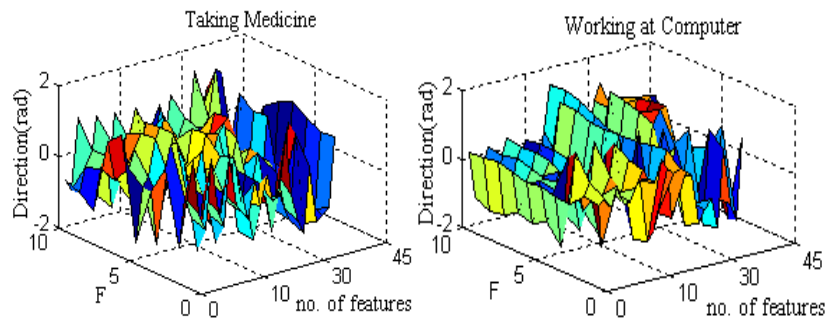


Fig. 6. Direction angles features of 15 joint points having different activities used in routine-log Real-AR system.

At last, by applying the spectral k-mean clustering to vector D (i.e. 1×60), the explicit symbolic structure can be obtained. More details about this process can be found in [34-37]. Nevertheless, the proposed method doesn't enquire the 3D subspace structure, only need to ensure that the 3D structures from overall features subspaces are preferred by proper discriminating among different classes. We applied such data for activity recognition purpose using HMMs. **Fig. 7** consists of following steps as; (1) it is started with continuous depth silhouettes as an input, (2) feature generation techniques are applied over skeleton models having joint points identification [38], (3) it used quantized algorithms to symbols of data

elements [39] of the joint body parts and (4) codebook are generated for training/recognition of each activity.

In-order to improve the real-time activity detection and recognition response time for codebook size, we set an optimal value of codebook size to get real-time system results.

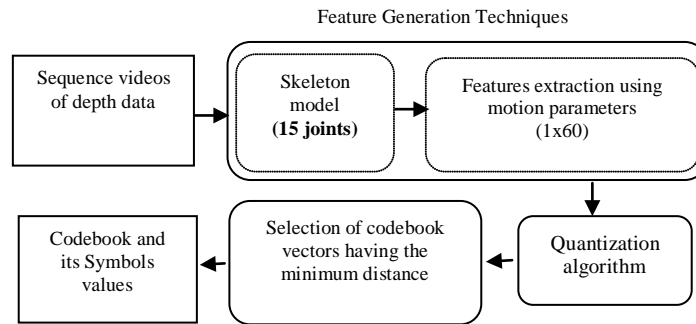


Fig. 7. Block diagram to analyze continuous video having joint points information, its features and symbol values.

2.3 Training and Recognition Using Recognizer Engine

Hidden Markov model (HMM) is considered as a generative probabilistic model that intends to describe the temporal sequences of uncertain input data. Then, it processes hidden states from observed data [40-42] based on its transition probabilities. There are four different probabilistic values used in HMM scenarios such as initial probability of the states, emission probability matrix, observation symbols probability and transition probability matrix. During training HMM, extracted features of each activity is used to train the four-states-left-to-right HMM model [43-45] and all activities are determined their specific likelihood values. In Fig. 8, all four states are used to measure likelihood values based on probabilistic values. Its values are embedded in-between 0-1.

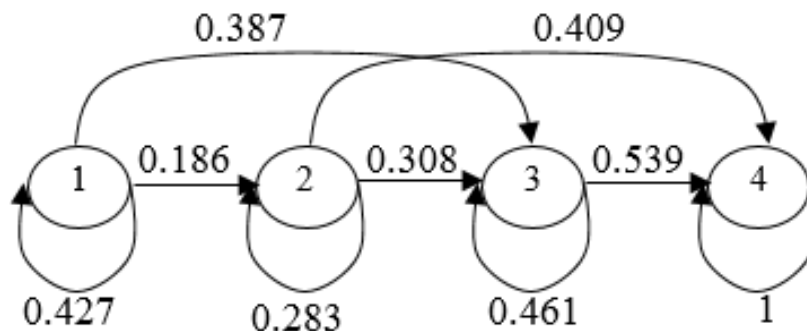


Fig. 8. Hand Clapping HMM structure and transition probabilities after training.

The feature vectors represented as symbol sequences are acquired to recognize an activity from the codebook creation and symbol assortment. It is applied on all of the trained HMMs to determine the likelihood [46] and then one among all is selected having the highest probability [47] to identify an activity. Thus, to test a feature vector sequence O , the HMMs act as

$$Act_{Recog} = \underset{i=1,2,\dots,N}{arg\ max} \{P_r(O | HMM_i)\}, \quad (3)$$

where Act_{Recog} is based on likelihood [48] of O on corresponding trained activity HMM_i .

3. Experimental Setting and Results

We make the quantitative evaluation of the proposed approach for accuracy evaluation on two annotated datasets against two different baseline methods. In experimental environment, a depth camera is adjusted at the top corner of the room. Each subject moves in front of camera inside the room [49] having a distance range of 2.5-2.8 meter [50]. The depth camera captures different activities from both the front and side views of the subjects.

In-order to train the system, we built the depth silhouette database including six most common smart home activities such as *walking*, *exercise*, *cooking*, *taking medicine*, *hand clapping* and *working at computer* performed by seven different volunteers. Each activity is represented in a video clip of 20 frames. We used 55 clips from each activity to build the training data and training data contained 6,600 activity depth images. In testing, all six activities are randomly performed and recognized. Finally, we produce different routine logging files to improve the quality of life.

3.1 System Interfaces

Meanwhile, we evaluate the performance of the proposed approach by reconstructing 3D skeleton motion from real-world 3D projection stemming from depth pointer-trackers of videos in training phase as presented in Fig. 9.

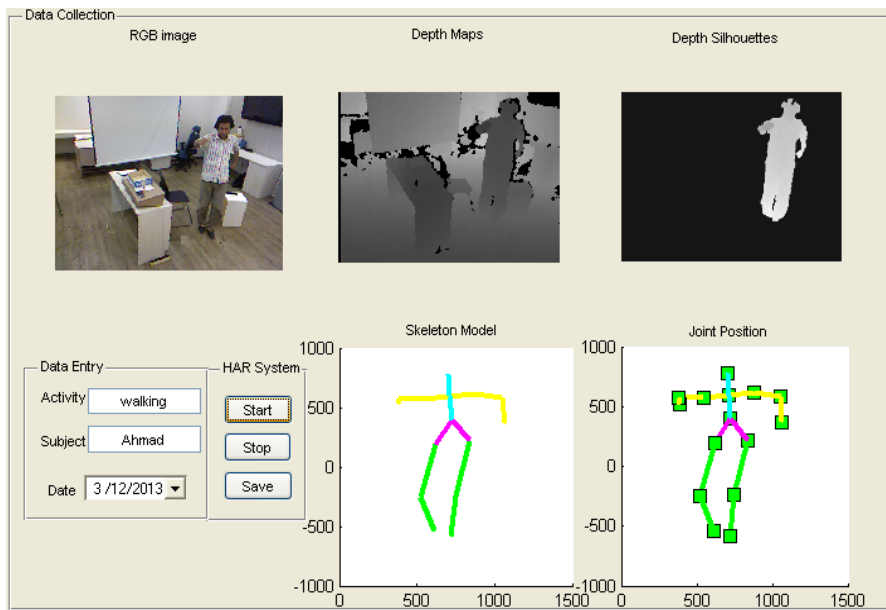


Fig. 9. Training data collection interface of proposed real-time routine logging system.

While, to recognize the desired activities, the information of the joint points is made available to be used to calculate the motion parameters attributes and incorporated the trained HMMs as depicted in Fig. 10.

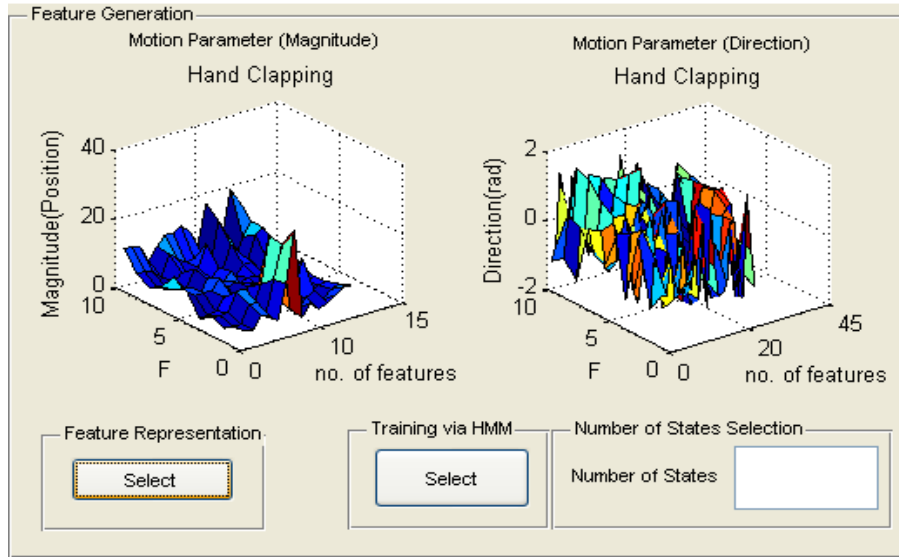


Fig. 10. Testing interface including feature generation and trained HMM toolbox for proposed real-time routine logging system.

3.2 Recognition Analysis of Continuous Routine Logging Systems

The depth cameras that observe the daily life actions of various subjects over definite hours in a day examine continuous video. Based on these measurements, we explore the performance of the two approaches such as conventional and proposed approaches.

The linear discriminant analysis (LDA) on principle component (PC) features is used as a conventional approach. The recognition comparisons are performed between originally recognized activities and the ground truth data using depth silhouettes (see Fig. 11). The PC extracted global feature properties are followed by the implementation of LDA (i.e., LDA on PC features) [12] to examine the behavior of each activity mentioned. It can be noticed that some activities are not identified accurately like *cooking*, *exercise*, *taking medicine* and *hand clapping* which ultimately caused low recognition rate.

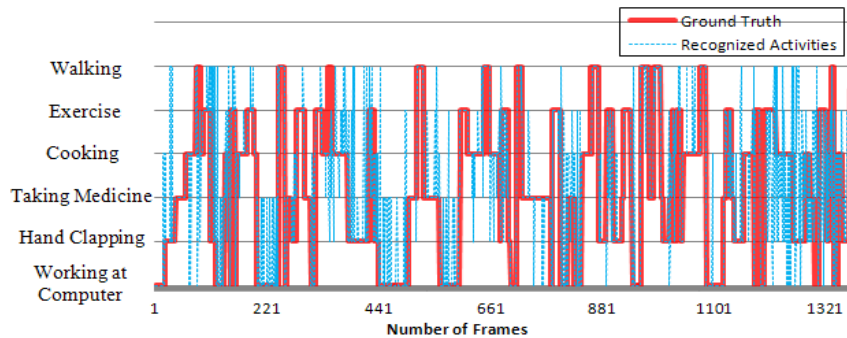


Fig. 11. Real time recognition of human activities using the LDA on PC feature-based approach.

However, the recognition outcomes of the proposed routine logging approach explores more persistently matching between the recognized activities and the ground truth and is illustrated in **Fig. 12**. Hence, it is proved that the proposed approach achieved higher recognition rate against the conventional approach as shown in **Table 1**.

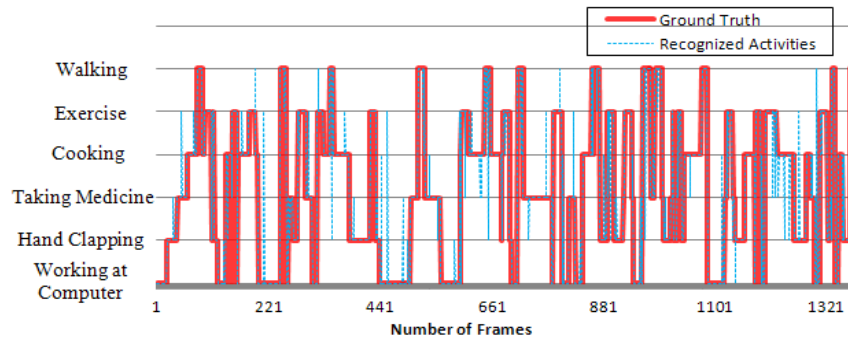


Fig. 12. Real time recognition of human activities using routine logging system based on the proposed routine logging feature-based approach.

3.3 Comparison of Recognition Results of the Conventional and Proposed Routine Logging Approaches via Depth Images

In this section, a comparison is made between the proposed routine logging system and the conventional approach by considering the depth silhouettes. The two routine logging systems are implemented to study relatively. Out of two systems, one covers LDA on PC features while the second system includes the motion factors of the body joints for Real-AR. For fair comparison, pre/post processing stages that includes human silhouettes extraction, training and running patterns of logging system remain unchanged. **Fig. 13** shows the assessments of the recognition results of both LDA on PC attributes and the proposed routine logging factors. Hence, the proposed approach improved the recognition rate for almost all the smart home activities.

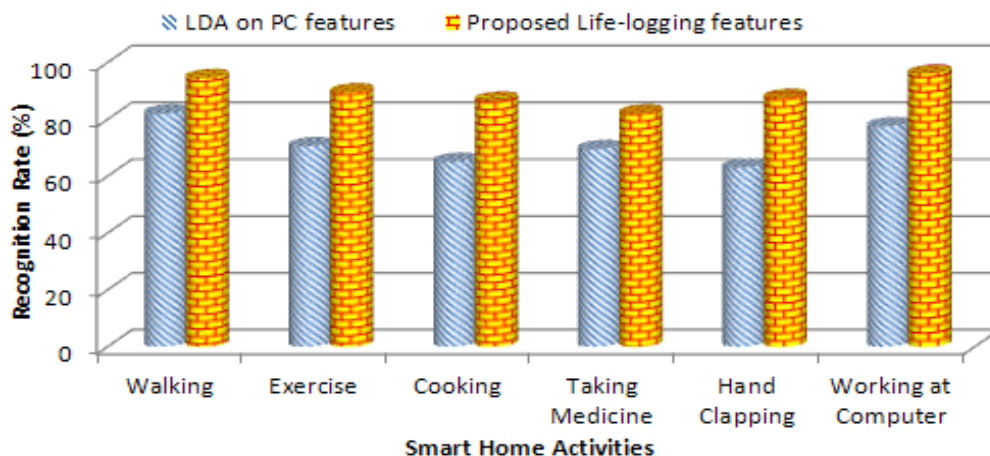


Fig. 13. Recognition rate comparison between conventional and proposed approach using depth images.

By considering **Table 1**, we calculated the recognition accuracy between existing and proposed methods. In case of existing as LDA on PC, Rf and ICA features, mean recognition rate are 71.25%, 72.83% and 74.58%, respectively. On the other hand, the proposed routine-logging approach show the improved recognition rate of 89.33% as compare to LDA on PC features via depth images. The proposed system operated about 18 frames per second during real-time scenarios.

Table 1. Recognition results of both conventional and proposed features approaches on depth images

Activities	LDA on PC Features	Rf transform features [51]	ICA Features [52]	Proposed Routine- logging Features
Walking	82.0	83.50	84.50	94.50
Exercise	70.50	70.0	74.50	89.50
Cooking	65.0	68.50	68.0	86.50
Taking Medicine	69.50	72.50	74.0	82.0
Hand Clapping	63.0	67.0	69.50	87.50
Working at Computer	77.50	75.50	77.0	96.0
Mean Recognition Rate	71.25	72.83	74.58	89.33

Finally, the overall comparison show that proposed routine logging approach provide better features and higher recognition rate than the LDA on PCA features, respectively. All experiments were implemented on 3.4GHz Inter Core 2 Duo Processor having 16GB RAM and a depth camera.

3.4 Experiments with MSRAction3D dataset

There is a benchmark dataset in the field of action recognition which is known as MSRAction3D dataset. It consists of twenty different actions captured by RGB-D camera. We performed a cross subject test and compared its results with other state-of-the-arts algorithms to evaluate the proposed system model (see **Table 2**).

Table 2. Comparison between our and conventional algorithms using MSRAction3D dataset

Methods	Recognition rate (%)
Bag of 3D points [53]	74.7
Eigenjoints [54]	82.3
STOP features [55]	84.8
Proposed algorithm	89.6

5. Conclusion

We proposed a real-time routine logging system to recover complex 3D human motions from multi-camera retrieved images based on matured feature representation processes. In addition, the proposed invariant scaled and translated skeletal model empirically proves to be more effective to represent complex human motion. Once the system is well trained, it provides maximum likelihood-based activity data and generates routine logs to improve life routine. Relying on the outcomes of the experiments, improved recognition rate of 89.33% is attained against the conventional approach as 71.25%. Moreover, the conventional approach has lack of confidence in serious body torso rotation and dealing with single depth features

attribute, which causes low recognition rates.

The proposed approach can be easily deployed to monitor and control smart home/office/hospital environments in a daily routine life.

References

- [1] S. Khawandi, B. Daya, and P. Chauvet, "Integrated monitoring system for fall detection in elderly," in *Proc. of IEEE Conference on Audio, Language and Image Processing (ICALIP)*, pp. 62-67, July 16-18, 2012. [Article \(CrossRef Link\)](#).
- [2] X. Zhang, T. Huang, Y. Tian, and W. Gao, "Background-modeling-based adaptive prediction for surveillance video coding," *IEEE Transactions on Image Processing*, vol. 23, no. 2, pp. 769-784, February, 2013. [Article \(CrossRef Link\)](#).
- [3] A. Jalal and M. A. Zeb, "Security enhancement for e-learning portal," *International Journal of Computer Science and Network Security*, vol. 8, pp. 41-45, 2008. [Article \(CrossRef Link\)](#).
- [4] J. Jo, S. Lee, and J. W. Kim, "Software-defined home networking devices for multi-home visual sharing," *IEEE Transactions on Consumer Electronics*, vol. 60, no. 3, pp. 534-539, March, 2014. [Article \(CrossRef Link\)](#).
- [5] A. Jalal and S. Kim, "The Mechanism of Edge Detection using the Block Matching Criteria for the Motion Estimation," *Human Computer Interaction*, pp.484-489, January, 2005. [Article \(CrossRef Link\)](#).
- [6] A. Jalal and I. Uddin, "Security architecture for third generation 3G using GMHS cellular network," in *Proc. of IEEE Conference on Emerging Technologies*, pp. 74-79, November 12-13, 2007. [Article \(CrossRef Link\)](#).
- [7] Y. He and Y. Li, "Physical activity recognition utilizing the built-in kinematic sensors of a smartphone," *International Journal of Distributed Sensor Networks*, vol. 2013, pp. 1-10, 2013. [Article \(CrossRef Link\)](#).
- [8] A. Jalal and Y. Rasheed, "Collaboration achievement along with performance maintenance in video streaming," in *Proc. of IEEE Conference on Interactive Computer Aided Learning*, pp. 1-8, September 26-28, 2007. [Article \(CrossRef Link\)](#).
- [9] A. Jalal, M.Z. Uddin, J. T. Kim, and T. Kim, "Daily Human Activity Recognition Using Depth Silhouettes and R Transformation for Smart Home," in *Proc. of International Conference on Smart Homes and Health Telematics*, pp. 25-32, 20-22 June, 2011. [Article \(CrossRef Link\)](#).
- [10] A. Jalal, S. Kamal and D. Kim, "Facial expression recognition using 1D transform features and Hidden Markov model," *Journal of Electrical Engineering and Technology*, vol. 12, no. 4, pp. 1657-1662, 2017. [Article \(CrossRef Link\)](#).
- [11] E. Kim, S. Helal and D. Cook, "Human activity recognition and pattern discovery," *IEEE Pervasive Computing*, vol. 9, no. 1, pp. 48-53, January, 2010. [Article \(CrossRef Link\)](#).
- [12] A. Jalal, M. Z. Uddin, and T.-S. Kim, "Depth video-based human activity recognition system using translation and scaling invariant features for life logging at smart home," *IEEE Transactions on Consumer Electronics*, vol. 58, no. 3, pp. 863-871, March, 2012. [Article \(CrossRef Link\)](#).
- [13] A. Jalal and S. Kim, "Algorithmic implementation and efficiency maintenance of real-time environment using low-bitrate wireless communication," in *Proc. of Proceedings IEEE International workshop on software technologies for future embedded and ubiquitous systems*, 2006. [Article \(CrossRef Link\)](#).
- [14] A. Jalal, S. Kamal, and D. Kim, "A depth video sensor-based life-logging human activity recognition system for elderly care in smart indoor environments," *Sensors*, vol. 14, no. 7, pp. 11735-11759, July, 2014. [Article \(CrossRef Link\)](#).

- [15] A. Farooq, A. Jalal and S. Kamal, "Dense RGB-D Map-Based Human Tracking and Activity Recognition using Skin Joints Features and Self-Organizing Map," *KSII Transactions on Internet and Information Systems*, vol. 9, no. 5, pp. 1856-1869, May, 2015. [Article \(CrossRef Link\)](#).
- [16] A. Jalal, "IM-DailyDepthActivity dataset," *Intelligent Media Lab*, 2017, [Online; accessed 14 April- 2017]. [Article \(CrossRef Link\)](#).
- [17] A. Jalal, M. Z. Uddin, J. T. Kim, and T.-S. Kim, "Recognition of human home activities via depth silhouettes and R transformation for smart homes," *Indoor and Built Environment*, vol. 21, no. 1, pp. 184-190, September, 2011. [Article \(CrossRef Link\)](#).
- [18] A. Jalal, J. T. Kim, and T.-S. Kim, "Human activity recognition using the labeled depth body parts information of depth silhouettes," *International Symposium on Sustainable Healthy Buildings*, pp. 1-8, 27-29 February, 2012. [Article \(CrossRef Link\)](#).
- [19] W. Lin, M.-T. Sun, R. Poovandran, and Z. Zhang, "Human activity recognition for video surveillance," in *Proc. of IEEE International Symposium on Circuits and Systems*, pp. 2737-2740, 18-21 May, 2008. [Article \(CrossRef Link\)](#).
- [20] A. Jalal, Y. Kim, S. Kamal, A. Farooq and D. Kim, "Human daily activity recognition with joints plus body features representation using Kinect sensor," in *Proc. of IEEE Conference on Informatics, Electronics and Vision*, pp. 1-6, 15-18 June, 2015. [Article \(CrossRef Link\)](#).
- [21] A. Jalal and Y. Kim, "Dense depth maps-based human pose tracking and recognition in dynamic scenes using ridge data," in *Proc. of IEEE Conference on Advanced Video and Signal-based Surveillance*, pp. 119-124, 26-29 August, 2014. [Article \(CrossRef Link\)](#).
- [22] A. Jalal, S. Kamal, and D. Kim, "Shape and motion features approach for activity tracking and recognition from Kinect video camera," in *Proc. of IEEE Conference on Advanced Information Networking and Applications*, pp. 445-450, 24-27 March, 2015. [Article \(CrossRef Link\)](#).
- [23] L. Xia, C.-C. Chen, and J. K. Aggarwal, "View invariant human action recognition using histograms of 3D joints," in *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 20-27, 16-21 June, 2012. [Article \(CrossRef Link\)](#).
- [24] O. Oreifej and Z. Liu, "HON4D: Histogram of oriented 4D normals for activity recognition from depth sequences," in *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 716-723, 23-28 June, 2013. [Article \(CrossRef Link\)](#).
- [25] S. Kamal, A. Jalal and D. Kim, "Depth Images-based Human Detection, Tracking and Activity Recognition Using Spatiotemporal Features and Modified HMM," *Journal of Electrical Engineering and Technology*, pp. 1921-1926, 2016. [Article \(CrossRef Link\)](#).
- [26] Y. Ming, G. Wang and C. Fan, "Uniform Local Binary Pattern Based Texture-Edge Feature for 3D Human Behavior Recognition," *Plos One*, vol. 10, no. 5, pp. 1-15, 2015. [Article \(CrossRef Link\)](#).
- [27] K. Charalampous and A. Gasteratos, "On-line deep learning method for action recognition," *Pattern analysis and applications*, vol. 19, no. 2, pp. 337-354, 2016. [Article \(CrossRef Link\)](#).
- [28] A. Jalal and S. Kim, "Global security using human face understanding under vision ubiquitous architecture system," *World Academy of Science, Engineering, and Technology*, vol. 13, pp. 7-11, 2006. [Article \(CrossRef Link\)](#).
- [29] L. Piyathilaka and S. Kodagoda, "Gaussian mixture based HMM for human daily activity recognition using 3D skeleton features," in *Proc. of IEEE Conference on Industrial Electronics and Applications*, pp. 567-572, 19-21 June, 2013. [Article \(CrossRef Link\)](#).
- [30] A. Jalal, Y. Kim, and D. Kim, "Ridge body parts features for human pose estimation and recognition from RGB-D video data," in *Proc. of IEEE Conference on Computing, Communication and Networking Technologies*, pp. 1-6, 11-13 July, 2014. [Article \(CrossRef Link\)](#).
- [31] N.-G. Cho, Y.-J. Kim, U. Park, J.-S. Park and S.-W. Lee, "Group activity recognition with group interaction zone based on relative distance between human objects," *International Journal of Pattern Recognition and Artificial intelligence*, vol. 29, no. 5, pp. 1-15, August, 2015. [Article \(CrossRef Link\)](#).

- [32] A. Jalal and M. A. Zeb, "Security and QoS Optimization for distributed real time environment," in *Proc. of IEEE Conference on Computer and Information Technology*, pp. 369-374, 16-19 October, 2007. [Article \(CrossRef Link\)](#).
- [33] A. Jalal and A. Shahzad, "Multiple facial feature detection using vertex modeling structure," in *Proc. of IEEE Conference on Interactive Computer Aided Learning*, pp. 1-7, 26-28 September, 2007. [Article \(CrossRef Link\)](#).
- [34] A. Jalal, S. Lee, J. Kim, and T. Kim, "Human activity recognition via the features of labeled depth body parts," in *Proceedings Smart Homes Health Telematics*, pp. 246-249, 2012. [Article \(CrossRef Link\)](#).
- [35] A. Jalal, J. T. Kim, and T.-S Kim, "Development of a life logging system via depth imaging-based human activity recognition for smart homes," in *Proc. of Proceedings of the International Symposium on Sustainable Healthy Buildings*, pp. 91-95, 2012. [Article \(CrossRef Link\)](#).
- [36] A. Jalal, S. Kamal and D. Kim, "Individual Detection-Tracking-Recognition using depth activity images," in *Proc. of IEEE Conference on Ubiquitous Robots and Ambient Intelligence*, pp. 450-455, 28-30 October, 2015. [Article \(CrossRef Link\)](#).
- [37] A. Jalal, S. Kim and B. J. Yun, "Assembled algorithm in the real-time H.263 codec for advanced performance," in *Proc. of IEEE Workshop on Enterprise Networking and Computing in Healthcare Industry*, pp.295-298, 23-25 June, 2005. [Article \(CrossRef Link\)](#).
- [38] A. Jalal, and S. Kim, "A complexity removal in the floating point and rate control phenomenon," *Korea Multimedia Society*, pp. 48-51, 2005. [Article \(CrossRef Link\)](#).
- [39] A. Jalal and S. Kim, S, "Advanced performance achievement using multi-algorithmic approach of video transcoder for low bit rate wireless communication," *International Journal on Graphics, Vision and Image Processing*, vol. 5, no. 9, pp. 27-32, September 2005. [Article \(CrossRef Link\)](#).
- [40] M. Rossi, S. Benatti, E. Farella and L. Benini, "Hybrid EMG classifier based on HMM and SVM for hand gesture recognition in prosthetics," in *Proc. of IEEE Conference on Industrial Technology*, pp. 1700-1705, 17-19 March, 2015. [Article \(CrossRef Link\)](#).
- [41] A. Jalal, N. Sharif, J. T. Kim, and T.-S. Kim, "Human activity recognition via recognized body parts of human depth silhouettes for residents monitoring services at smart homes," *Indoor and Built Environment*, vol. 22, pp. 271-279, 2013. [Article \(CrossRef Link\)](#).
- [42] S. Kamal and A. Jalal, "A hybrid feature extraction approach for human detection, tracking and activity recognition using depth sensors," *Arabian Journal for Science and Engineering*, vol. 41, no. 3, pp. 1043-1051, March 2016. [Article \(CrossRef Link\)](#).
- [43] A. Jalal, S. Kamal, and D. Kim, "Depth map-based human activity tracking and recognition Using body joints features and self-organized map," in *Proc. of IEEE Conference on Computing, Communication and Networking Technologies*, pp. 1-6, 11-13 July, 2014. [Article \(CrossRef Link\)](#).
- [44] A. Jalal, S. Kamal and D. Kim, "Depth Silhouettes Context: A new robust feature for human tracking and activity recognition based on embedded HMMs," in *Proc. of IEEE 12th International Conference on Ubiquitous Robots and Ambient Intelligence*, pp. 294-299, 28-30 October, 2015. [Article \(CrossRef Link\)](#).
- [45] A. Jalal, S. Kamal, A. Farooq and D. Kim, "A spatiotemporal motion variation features extraction approach for human tracking and pose-based action recognition," in *Proc. of IEEE Conference on Informatics, Electronics and Vision*, pp. 1-6, 15-18 June, 2015. [Article \(CrossRef Link\)](#).
- [46] T. Kim, A. Jalal, H. Han, H. Jeon and J. Kim, "Real-Time Life Logging via Depth Imaging-based Human Activity Recognition towards Smart Homes Services," in *Proc. of Proceedings International Symposium on Renewable energy sources and healthy buildings*, pp. 63, 2013.
- [47] A. Jalal, Y.-H. Kim, Y.-J. Kim, S. Kamal and D. Kim, "Robust human activity recognition from depth video using spatiotemporal multi-fused features," *Pattern Recognition*, vol. 61, no. 1, pp. 295-308, January 2017. [Article \(CrossRef Link\)](#).

- [48] A. Kundu, Y. He and M.-Y. Chen, "Alternatives to variable duration HMM in handwriting recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 11, pp. 1275-1280, November 1998. [Article \(CrossRef Link\)](#).
- [49] A. Jalal, Shaharyar Kamal and Daijin Kim, "Human depth sensors-based activity recognition using spatiotemporal features and hidden markov model for smart environments," *Journal of computer networks and communications*, vol. 2016, pp. 1-11, 2016. [Article \(CrossRef Link\)](#).
- [50] A. Jalal and S. Kamal, "Real-Time Life Logging via a Depth Silhouette-based Human Activity Recognition System for Smart Home Services," in *Proc. of Proceedings of the IEEE International Conference on Advanced Video and Signal-based Surveillance*, pp. 74-80, 2014. [Article \(CrossRef Link\)](#).
- [51] J. M. Carmona and J. Climent, "Action recognition using Rf transform on Optical flow images," in *Proc. of International Joint Conference on Computer vision, imaging and computer graphics theory and Applications*, pp. 266-271, 27 February, 2017. [Article \(CrossRef Link\)](#).
- [52] D.-M. Tsai and W.-Y. Chiu, "A real-time ICA-based activity recognition in video sequences," in *Proc. of International Conference on Machine vision Applications*, pp. 467-470, 20-23 May, 2013. [Article \(CrossRef Link\)](#).
- [53] W. Li, Z. Zhang and Z. Liu, "Action recognition based on a bag of 3d points," in *Proc. of Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition*, pp. 9-14, 13-18 June, 2010. [Article \(CrossRef Link\)](#).
- [54] X. Yang and Y. Tian, "Eigenjoints-based action recognition using naive-bayes-nearest-neighbor," in *Proc. of Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition*, pp. 14-19, 16-21 June, 2012. [Article \(CrossRef Link\)](#).
- [55] A. Vieira, E. Nascimento, G. Oliveira, Z. Liu, M. Campos, Stop: Space-time occupancy patterns for 3d action recognition from depth map sequences, in *Proc. of Proceedings of the Iberoamerican Congress on Pattern Recognition*, pp. 252-259, 3-6 September, 2012. [Article \(CrossRef Link\)](#).



Ahmad Jalal received his B.S. degree in Computer Science from Iqra University, Peshawar, Pakistan and M.S. degree in Computer Science from Kyungpook National University, Republic of Korea. He received his Ph.D. degree in the Department of Biomedical Engineering at Kyung Hee University, Republic of Korea. He becomes “Research Associate” in Pohang University of Science and Technology (POSTECH). His research interest includes human computer interaction, image processing, and computer vision.



Shaharyar Kamal received his B.S. degree in Software Engineering from City University of Science and Information Technology, Peshawar, Pakistan and M.S. degree in Computer Engineering from Mid Sweden University, Sweden. He is currently enrolled as Ph.D. candidate in the Department of Radio and Electronics Engineering at Kyung Hee University, Republic of Korea. His research interest includes advanced wireless communication, image and signal processing.



Dong-Seong Kim received his Ph. D degree e in EECS, Seoul National University, Seoul, Korea. From 1994 to 1998, he worked as a full-time Researcher in ERC-ACI at Seoul National University. From September 2000 to December 2001, he worked as a part-time Lecturer in department of information and communication at Dong-Guk University. He is currently Professor in Department of IT Convergence Engineering, School of Electronic Engineering, Kumoh National Institute of Technology, Gumi, Korea. He has other responsibilities as Director and Head of Convergence Technology Institute, ICT-CRC, Korea. His main research interests include Industrial networked control system, Fieldbus and Real-time systems and wired/wireless military networks.