2017 IEEE International Conference on Big Knowledge

Discovering Knowledge by Behavioral Analytics for Elderly Care

Bruce X.B. Yu, Keith C.C. Chan
Department of Computing
The Hong Kong Polytechic University
Kowloon, Hong Kong
bruce.xb.yu@connect.polyu.edu.hk, keith.chan@polyu.edu.hk

***Abstract*— Population aging is a phenomenon affecting many developed and developing countries in the world and many of them are expecting to face various social and economic problems as a result. In response, there have been effort to improve elderly care. Particularly, preventing dementia and taking better care of dementia patients are considered very important. In this paper, we present a system we developed to recognize indoor daily routines of elderly people so that their needs and interests can be better served. Our system uses a Kinect v2 sensor network covering the whole indoor living area of an elderly. Daily routine is the highest level of activity recognition derived based on the duration and complexity of the movement captured in a Kinect network. The relationship between movement and activities are discovered using both features extracted from sensor data and a machine learning method called AdaBoost. Based on the activities discovered, the system can be useful for assisting elderly people to better live their lives with automatic reminders, alarm services related to safety and security, the feeding of healthy information, and even in the discovering of dementia related symptoms. With the Kinect v2 sensor network, the proposed system will work even under poorly illuminated conditions and even if the elderlies do not use any wearable sensors.**

Keywords-Knowledge Discovery; Activity Recognition

I. INTRODUCTION

Population aging is an issue of major concern to many different countries all around the world. Statistics from the Global Age Watch index indicates that the population aged 60 or over will reach approximately 21.5% (2,1 billion) of the total population of the world by 2050 [1]. Elderly people have higher risk to be affected by degenerative diseases from cognitive diseases to functional impairment. One of the most common of all is dementia. About 5% of the elderly population aged over 65 are diagnosed with it [2]. The total number of such elderlies stands currently at 47.5 million and the figure is expected to be tripled by 2050 [3]. The increase in spending related to providing healthcare for the elderlies is expected to continue to increase. Particularly, the costs for taking care of dementia patients is estimated to be about US$818 billion, which is projected to be rising to US$2 trillion by 2030 [4]. Funding for dementia-related research has therefore also been increasing from US$ 650 million to 753 million [5] in the most recent three years.

Dementia is a generic term that describes chronic or progressive dysfunction of cortical and subcortical function that results in complex cognitive decline. The symptoms of dementia include common cognitive changes in disturbances of mood, behavior, and personality [6]. Dementia diagnosis is mostly based on observing for symptoms in the hope that the diseases can be discovered earlier in cheaper, more reliable methods. Normally, symptoms must have appeared for at least six months for a subject to be considered to have dementia according to some cognitive test methods [7]. Hence, observation based diagnosis are usually not efficient enough to capture the progressive decline of mental or physical functioning unless they are monitored over a period. If diagnosis can be made earlier, patients could take preventative medicine and their family members could prepare for a plan earlier.

Researchers have developed various staging scales to represent severity of dementia and among them, the Clinical Dementia Rating (CDR) is the most popular [8]. The CDR considers dementia to be a process that consists of seven stages based on severity of cognitive decline:

Stage 1: No Cognitive Decline
Stage 2: Very Mild Cognitive Decline
Stage 3: Mild Cognitive Decline
Stage 4: Moderate Cognitive Decline
Stage 5: Moderately Severe Cognitive Decline
Stage 6: Severe Cognitive Decline
Stage 7: Very Severe Cognitive Decline

In the last three stages, a dementia patient would require one-to-one assistance by family members or caregivers to live his or her daily life. Fortunately, dementia can be discovered in the first three stages. Main symptoms that are noticeable are the deterioration of mental health, memorization and physical performance decline. For instance, in the third stage, which is also known as Mild Cognitive Impairment (MCI) stage, the symptoms that become noticeable to family members include:

a) Increased forgetfulness,
b) Slight difficulty concentrating,
c) Decreased work performance,
d) Get lost more often, and
e) Having difficulty to find the right words.

It is worth mentioning that the average duration of the third stage (a.k.a. MCI) is seven years, which is a relatively long dysfunction process [9]. It is at this stage where early prevention and preparation could have effective benefit. Early diagnosis could alleviate significant loss of Quality of Life (QoL) for the patient and his/her family members. What we can do is to discover early symptoms of dementia at

978-1-5386-3120-1/17 $31.00 © 2017 IEEE
DOI 10.1109/ICBK.2017.18

earlier stages to delay and prevent the process of deterioration. Towards this goal, we propose an activity recognition system focusing on symptoms in the third stage of dementia. The system will first detect for various "activities of daily living" (ADLs) of an elderly person so that daily routines can be analyzed based on features extracted from sensor data and the machine learning method called AdaBoost to discover if there is any sign of early development of dementia symptoms.

In section II, we will review the related works of activity recognition. Section III describes the proposed activity recognition methods and daily routine analyzing algorithm, follow by a system implementation details in section IV and the evaluation in section V. Section VI provides discussion and conclusion.

II. RELATED WORKS

A. Activity Level and Activity Recognition

Many Ambient Assistive Living (AAL) technologies utilizing various sensors have been developed to support elderlies to live independently. Among these technologies, human activity recognition (HAR) is an essential component of AAL systems. HAR gained increasing industrial and academic attention due to its various potential applications not only in elderly care but also in other domains like entertainment (interactive game) and military (tactical decision making) [10]. HAR can be roughly divided into external and wearable sensors. The former refers to where sensors are installed in predetermined places for residents to interact with while the latter attaches sensors to a user. According to a survey on AAL tools for elderly adults, fig. 1 shows that common methods of HAR have three main categories namely wearable (Activity Recognition) AR, ambient AR and vision-based AR [11].

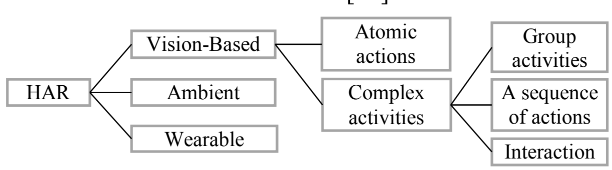

Fig. 1. Hierarchical structure of HAR healthcare systems

Our proposed solution for daily activity recognition uses vision-based approach as it is less intrusive for elderly users comparing with wearable sensors and easier to install in contrasting to ambient sensors. The vision-based AR is motivated by the recent development of cost-effective depth sensors which generates the 3D structure of an image where each pixel provides the distance between the real-world points to the sensor [12].

For better introducing the state of the art of vision based activity recognition, the complexity levels of the activities need to be defined and classified. Chaaraoui et al. [13] classified the vision based human behavior analysis according to the degree of semantics (motion, action, activity and behavior) and the time span (from frames to days). Table I shows activity level definition of five categories based on the approach utilized by Chaaraoui.

TABLE I. ACTIVITY LEVELS

Complexity	Duration	Examples
Human pose/gesture [14]	Frames, seconds	Appearance, hand gestures
Human action [15] [16] [17]	Seconds, minutes	Fall down, stand up, sit down
Human activity [18] [19]	Minutes, hours	Reading, cooking, sleeping
Human behavior [20] [21]	Hours, Days, Weeks	Morning routine, daily routine
Group activity interaction [22] [23]	Seconds to hours	Merging, planning, cooperation

The lowest level human pose or gesture estimation aims to achieve an accurate and fast joints approximation. Roy et al. [14] proposed a method to improve human detection accuracy by alleviating the black hole noise, which could further improve the HAR accuracy. Fall is the most researched action as it is the most probable cause for elderly people's death [15] [16]. Ben et al. [17] developed a Kinect wireless sensor network covering the whole living area, but the actions recognized are not relevant with elderly care. Comparing with action recognition, human activity has a longer time span from minutes to hours. Reddy et al. [18] and Luo [19] used similar approaches (MSR 3Ddaily dataset and SVM classifier) to recognize activities in the dataset with the latter having more activities but lower accuracy. The human behavior level is higher than the human activity level since it covers various daily activities like cooking, sleeping and watching TV etc. Kary [20] proposed a framework for predicting the human's location in a kitchen as a morning routine. Dell' Acqua et al. [21] developed a human activity recognition and monitoring system called HAREM to infer daily routines of elderlies. However, it just monitors seven activities (walk, jump, grab, stand, sit, lie, move up and done) and the Kinect network mentioned is just part of their future jobs. The most complex activity level is group activity interaction since it requires user identification. Elangovan [22] classified the group activities into human-vehicle, human-human and human-object and implemented the human-vehicle interaction through a Modified Sequential-Hidden Markov Model. Yang [23] proposed a method to detect abnormal group activities like fighting and chasing by calculating the action energy in human targets' occupancy area (head centered circle with an 80cm long radius).

B. Available Video Sensor for Activity Recognition

Widely used depth camera sensors are Time-of-Flight (ToF) cameras and structured light cameras like Microsoft Kinect for windows (K4W). Comparing with traditional stereo approach which generates the 3D structure from two or more viewpoints [24], depth cameras turn out to be more robust to illumination changes or shadows. Fig. 2 shows three types of depth sensors. Although HAR has been being researched since 1980s, the vision-based HAR experiences tremendous progress just in recent years with the release of low-cost K4W depth sensor. Kinect HAR datasets were emerging rapidly, which covers gaming actions, atomic actions, daily activities, arm gestures, and even hand gestures [12]. K4W shows the potential capability to be implemented in HAR system with its popularity, free SDK and low cost [15]. Given the K4W sensor, developing robust HAR systems needs to tackle with the issue of holes where depth information is undefined or occluded [25]. As shown in fig. 2 (c), the K4W v2 sensor is released by Microsoft on July 15, 2014, which provides improved capability from depth sensing to field of view, and from skeletal tracking to active Infrared (IR) [26]. The new version of K4W can acquire six body skeletons simultaneously with each skeleton has 25 joints [27].

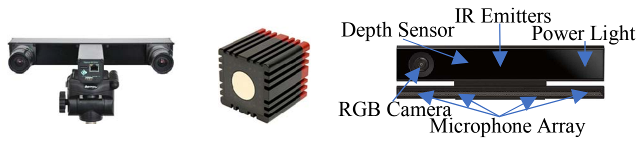

Fig. 2. Different types of stereo cameras

The state of the art literatures all try to solve HAR tasks by using machine learning technology. However, some simple activities are easier to be hard coded than using machine learning tools. Technically, there should be a balance between hard coding and machine learning. In our proposed solution, we not only code the skeleton joint data but also use machine learning to recognize activities.

C. *Daily Routine Knowledge Discovery*

There are few researches attempting to mining daily routine data. The duration of publicly available datasets is usually too short to provide some significant changes in daily routines [28] [29]. [30] claimed to discover activity patterns, but the patterns are for activity recognition rather than mining daily routines. Low resolution activities like waking, standing, lunch, sitting are not informational enough to discover reliable dementia symptoms. [31] proposed method to detect detailed cooking actions (mix, place, chop, pour, spoon scoop etc.) in kitchen, but the recognition resolution is not high enough to discover dementia symptoms.

D. *Challenges and Issues*

The K4W takes the advantage of depth sensor to avoid privacy issue encountered by traditional approaches [16] [18]. However, developing indoor HAR systems is confronted with other aspects of challenges and issues.

1) Limitation of Kinect: Objects that too close to the Kinect senor (closer than 20 inches) will cause black holes. The IR dots will be displaced farther in a mirror since the mirror reflect the dots from a symmetrical virtual space in the mirror [32]. To avoid occlusion, more than one calibrated sensors are required to view the environment from multiple directions. Maimone et al. introduced an automatic color calibration method to generate a real time stereo scene [33].

2) Activity Diversity and Accuracy: Some activities share similar features and confuse the system. The system developed by Luo [19] achieved a high average recognition accuracy, but the recognition accuracy of actions like reading book, writing and using laptop seem unacceptable. Some of the state of the art literatures list every feasible combination of selected features [18] and utilized classifier algorithms [34] to figure out the best solution for specific action recognition.

3) Scalability and Reusability: The large diversities of subject appearance make it difficult to find or make a suitable training datasets [25]. The emerging number of datasets for action recognition is still limited and need novel approaches to handle real world scenarios. In this case, current vision-based action recognition systems are faced with the challenge of adapting to multiple subjects [35].

III. PROPOSED METHOD

The SDK of K4W use "Body" class to represent skeleton joints. All joints in an acquired "Body" frame can be represented by different space structures namely: camera Space Point, color space point and depth space Point. Left part of fig. 3 shows a joint location's X, Y and Z coordinates in the camera space where the point (0, 0, 0) is the center of K4W's IR emitters, coordinate X grows to the sensors left, Y grows up and Z grows out to where the sensor is facing. Unlike in camera space, the X and Y component of a joint in the depth space signals the 2D location of the depth image. In right of fig. 3, 2D point (0, 0) is the top left corner of the depth image where the parallelogram corresponds to the acquired depth image. Besides, the Z coordinate is the same with that of camera space. In the remaining of this section, algorithms for action recognition will be explained in detail.

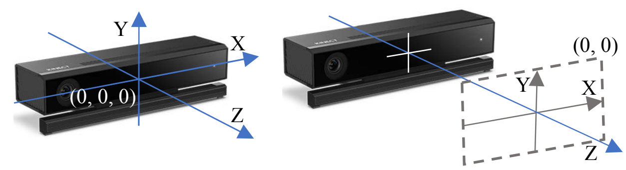

Fig. 3. 3D Location of camera space (left) and depth space (right)

A. *Body Joint Position Based*

1) Cross Door: Recognizing cross a door is firstly coordinate the "Body" joints which is structured in camera space to depth space. Then the area of the door is defined as a circle area by the SpineShoulder and SpineBase joints of "Body" where the SpineShoulder joint is the center while the distance between SpineShoulder and SpineBase (marked as DRadius in fig. 4) is the radius of the circle. The distance between the detected SpineShoulder of a subject and the door center in the coordinate Z direction is sued to determine if the user have crossed the door or not.

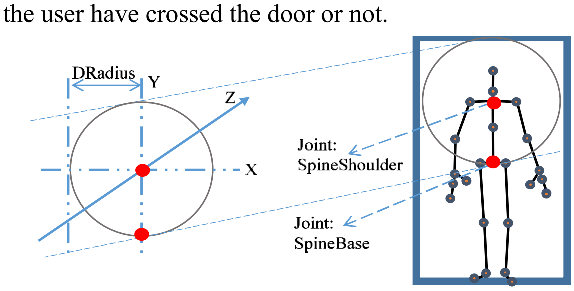

Fig. 4. Define SpineShoulder as door position

Another case treated as cross a door is when a Body loses tracking in the door area, which is the case that a door is not facing very straightly to the K4W sensor or a user enter a door and turn left or right immediately.

2) Fall Down and Sleep: When a user fall, all his joints should be close to the floor plane. Empirically, the most noticeable features for falling detection are Head and SpineMid joints. Floor plane corresponding to the camera space, which is refreshed and contained in each "Body" frame. A general plane equation is in (1).

$$Ax + By + Cz + D = 0 \tag{1}$$

Assuming a "Body" joint point J (x_j, y_j, z_j) and normal vector (A, B, C) in (1) is a unit vector, then the projection d of point J on the normal vector (A, B, C) is in (2).

$$d = \frac{|\ Ax_j + By_j + Cz_j + D\ |}{\sqrt{A^2 + B^2 + C^2}} \tag{2}$$

After generating the distance to the floor valued d, a threshold of d is defined to judge the fall down action. It is noted that the floor clip plane may not always be detectable, which will lead to a zero vector. Similarly, two feet and the head are relatively noise free features for detecting the sleep s action. When all the distances of these three joints from the floor are within predefined heights, it would be probably that the elderly subject is sleeping and the system will record the his/her sleeping time.

3) Walking Related: When the elderly is walking around home, all the joints should be moving. But some of the joints may be invalid due to occlusion. The joint SpineShoulder is chosen as a strong feature for walking detection. When the movement speed of the user is surpassing the threshold speed of walking detection, the system will recognize the walking action. locations of a cabinet storing medical pills or a desk for eating meals are defined by the Head joint with a distance threshold for judging the subject's appearance. By defining the location where he or she puts his medical drugs and tracking his or her appearance to a location is a way to roughly detect whether he or she takes the medicine.

B. Machine Learning Method

As described above, algorithms in body joint position based activity recognition can recognize noticeable featured actions like walking, fall and cross a door etc. However, more complex actions require machine learning method. Main tuning procedure of machine learning from data tagging, feature extracting to training and detection are shown in fig. 5.

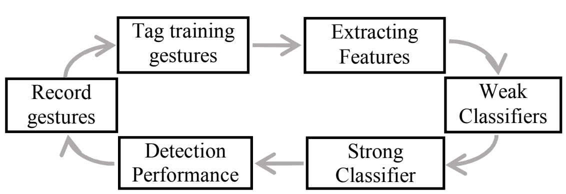

Fig. 5. Machine learning tuning procedure

We take drink action for introduction. Given 6 skeleton frames h_t ($h_t \in \mathbb{R}^{N\times(t+5)}$, N is the number of joints multiplied by joint data dimensions) of drink water, joints include HAND_RIGHT, WRIST_RIGHT, SPINE_SHOULDER, HEAD NECK, SHOULDER_RIGHT, and ELBOW_RIGHT will be used. Using the AdaBoost algorithm in [36], features listed in table II will extracted to form M weak threshold classifiers pool $G_i(h)$ as shown in table III, in which the hit is recorded as 0 and the miss is 1.

TABLE II. FEATURES FOR GESTURE RECOGNITION

Intra joint features	Inter joints features
1. Joints position distance 2. Angles between 3 joints 3. Velocity of angles 4. Acceleration of angles	1. Velocity in 3D space 2. Speed 3. Acceleration 4. Muscle force 5. Muscle torque

TABLE III. WEAK THRESHOLD CLASSIFIERS BASED ON FEATURES

	G_1	G_2	…	G_M
h_1	0	1	…	1
h_2	0	0	…	1
h_3	1	1	…	0
⋮	⋮	⋮		⋮
h_T	0	0	…	0

The classifier pool $G_i(h)$ will be firstly initialized with weights $w_i^{(m+1)}$(i = 1,2, …, M). Then all the weights will be optimized by gradient descent. In the m-th iteration of the gradient update loop, the weight $w_i^{(m+1)}$ will be updated by $\alpha_m = \ln((1 - err_m)/err_m)$ as $w_i^{(m)} e^{\pm\alpha_m}$, where err_m is calculated in (3).

$$err_m = \frac{\sum_{i=1}^{M} w_i e^{\alpha_m}}{\sum_{i=1}^{M} w_i} \tag{3}$$

The final strong classifier $G(h)$ is the sign function of the sum of the weak classifiers in (4).

$$G(h) = \operatorname{sign}\left(\sum_{m=1}^{M} \alpha_m\, G_m(h)\right) \tag{4}$$

IV. SYSTEM IMPLEMENTATION

In fig. 6, the system structure of the Kinect network for activity recognition contains three Kinect sensors hosted by three PCs with one of them being the Gateway. The network protocol used for the Kinect sensors is the TCP/IP socket through a Wi-Fi network. Each Kinect sensor is installed to a specific living area as shown in fig. 8. In that way, the whole living area of an elderly subject will be monitored except areas like toilet and warehouse. The time series data is stored in MongoDB for the ease of data retrieve and analysis. The server provides APIs to front-end applications of mobile devices or websites. The front-end applications are the portal for family members or caregivers to check the independent living reports containing daily routine status, the discovered abnormal symptoms. Meanwhile, some emergent scenarios will be detected and alarm other family members through e-mail or push a notification to mobile devices.

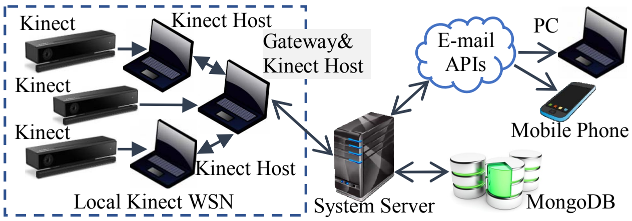

Fig. 6. Activity recognition system structure fors AAL

The workflow of activity recognition for an independent elderly adult is shown in fig. 7. Initially, the gateway of the system will start to wait for connection requests from Kinect hosts to set up connections. Complete connection, the gateway can control the Kinect hosts through web service. Before the Kinect hosts are ready for connection, each Kinect host needs to define locations like doors, objects and locations related to some actions like take pills or eat meals. After every Kinect host is ready for tracking, a user in the living area where the gateway may start the system by turning on the living room Kinect host. When the user leaves the current location through a predefined door, the system will send turn on commend to the Kinect host in the living area where he or she is just arrived based on the predefined door. In that way, the fundamental daily routine with an appearance level will be recorded. Building on the appearance level, other activities that occur in each living area can be recognized accordingly. Taking the activities in a bedroom for example, the system would record the elderly's sleeping hours. In case the elderly doesn't get up after ten o'clock in the morning, the system will send an email alarm to his or her family members.

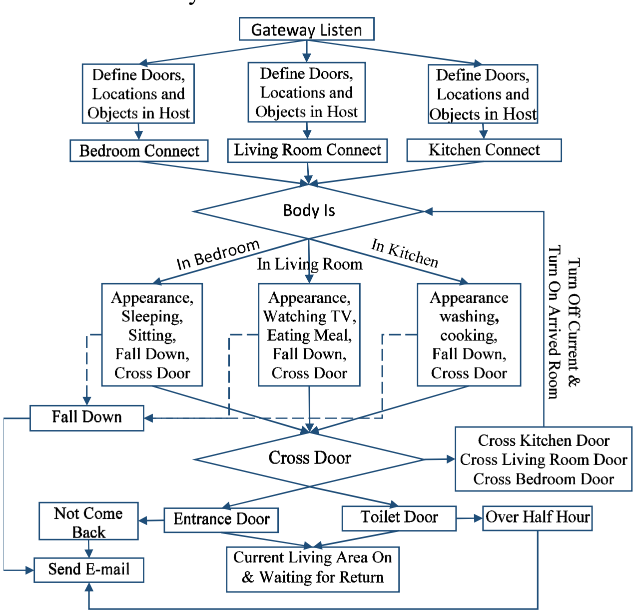

Fig. 7. Single person activity recognition system work flow

V. EXPERIMENTAL EVALUATION

To develop application using Kinect sensor, the developing environment need to meet the minimum requirements as stated by official website of Microsoft Kinect for Windows v2 [37]. This system is developed and tested under the developing environment as table IV shows.

TABLE IV. DEVELOPING AND TESTING ENVIRONMENT

Three PCs	Lenovo T440s i7, Lenovo 240x and ThinkCenter M93p.
Operation System	Window 10
Developing Environment	Visual Studio 2015 Enterprise
Developing Language	C# or C++

The system introduces a new attempt for indoor activity recognition, which monitors the entire living area and achieved a high level of activity recognition with relatively high resolution. Fig. 8 shows the testing environment in our intelligent home laboratory, which is a screen capture of the video demo of our HAR system. There are three Kinect sensors consisting the sensor network covering the bedroom, kitchen, and living room of indoor living environment.

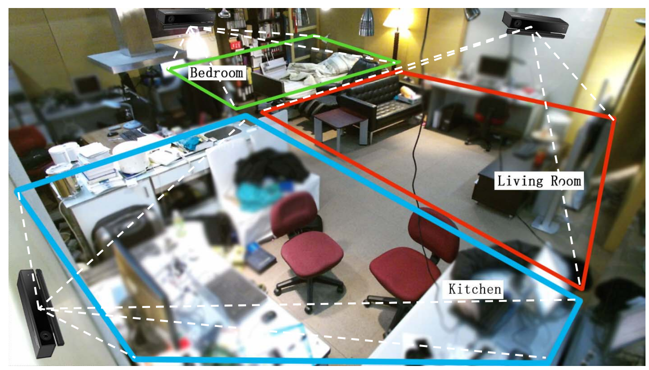

Fig. 8. Testing environment in the intelligent home

In terms of computational cost, table V shows the comparison of memory usage comparison on Lenovo T440s i7 during the HAR system gateway and Kinect host are working and when the Kinect studio is working alone. Analyzing the figures of the comparison table, it indicates that our HAR system runs normally without occupying much computational resources comparing with the performance of the Kinect studio provided by Microsoft.

TABLE V. CPU AND MEMORY USAGE COMPARISON

	Before Testing	**Kinect Studio**	**HAR Application**
CPU	3% 0.84GHz	64% 2.65GHz	20% 1.05GHz
Memory	3.0/7.9 GB (38%)	3.3/7.9 GB (42%)	3.1/7.9 GB (39%)

To develop application using K4W v2 sensor, the developing environment need to meet the minimum requirements as officially stated by Microsoft Kinect for Windows v2 [37]. This system is developed and tested under the developing environment as shown in table IV.

VI. CONLUSIONS

In this paper, we proposed a novel human activity recognition system for independent elderlies to alleviate the work load of care givers and family members. In an almost entire monitored and assistive indoor environment, elderlies may live independently with more confidence. The system uses both 3D joint features and machine learning to detect multiple levels of activities and capable to record daily routines. Provided the daily routines, abnormal dementia symptoms like memory and performance decline could be discovered by behavioral analytics.

The Kinect sensor network system could adapt to different household scenario usages by scaling to various numbers of rooms. Multiple features are used for activity recognition with good performance in terms of recognition and computational cost. However, there remains challenging issues to be solved. In the future, we will expand the system to multi-user activity recognition by using resident identification technologies. More robust and accurate activity recognition algorithm need to be developed. Last but not the least, frameworks using Kinect sensor network and other ambient assistive sensors are expected to emerge soon.

REFERENCES

[1] Global AgeWatch Index 2013: Insight report. 2013, HelpAge International.

[2] Umphred, D.A., et al., Neurological rehabilitation. 2013: Elsevier Health Sciences.

[3] Alzheimer's, A., 2015 Alzheimer's disease facts and fig.s. Alzheimer's & dementia: the journal of the Alzheimer's Association, 2015. 11(3): p. 332.

[4] Prince, M.J., World Alzheimer Report 2015: the global impact of dementia: an analysis of prevalence, incidence, cost and trends. 2015.

[5] Estimates of Funding for Various Research, Condition, and Disease Categories (RCDC). 2016.

[6] Lampit, A., H. Hallock, and M. Valenzuela, Computerized cognitive training in cognitively healthy older adults: a systematic review and meta-analysis of effect modifiers. PLoS Med, 2014. 11(11): p. e1001756.

[7] Lin, J.S., et al., Screening for cognitive impairment in older adults: a systematic review for the US Preventive Services Task Force. Annals of internal medicine, 2013. 159(9): p. 601-612.

[8] Reisberg, B., et al., The Global Deterioration Scale for assessment of primary degenerative dementia. The American journal of psychiatry, 1982.

[9] Seven Stages of Dementia | Symptoms & Progression. 2016; Available from: https://www.dementiacarecentral.com/aboutdementia/facts/stages/.

[10] Labrador, M.A. and O.D.L. Yejas, Human Activity Recognition: Using Wearable Sensors and Smartphones. 2013: CRC Press.

[11] Palumbo, F., et al., Sensor network infrastructure for a home care monitoring system. Sensors (Basel), 2014. 14(3): p. 3833-60.

[12] Aggarwal, J.K. and L. Xia, Human activity recognition from 3D data: A review. Pattern Recognition Letters, 2014.

[13] Chaaraoui, A.A., P. Climent-Pérez, and F. Flórez-Revuelta, A review on vision techniques applied to Human Behaviour Analysis for Ambient-Assisted Living. Expert Systems with Applications, 2012. 39(12): p. 10873-10888.

[14] Roy, S. and T. Chattopadhyay, View-Invariant Human Detection from RGB-D Data of Kinect Using Continuous Hidden Markov Model, in Human-Computer Interaction. Advanced Interaction Modalities and Techniques. 2014, Springer. p. 325-336.

[15] Parajuli, M., et al. Senior health monitoring using Kinect. in Communications and Electronics (ICCE), 2012 Fourth International Conference on. 2012. IEEE.

[16] Gasparrini, S., et al., A depth-based fall detection system using a Kinect(R) sensor. Sensors (Basel), 2014. 14(2): p. 2756-75.

[17] Ben Hadj Mohamed, A., et al. Using a Kinect WSN for home monitoring: principle, network and application evaluation. in Wireless Communications in Unusual and Confined Areas (ICWCUCA), 2012 International Conference on. 2012. IEEE.

[18] Reddy, V.R. and T. Chattopadhyay, Human Activity Recognition from Kinect Captured Data Using Stick Model, in Human-Computer Interaction. Advanced Interaction Modalities and Techniques. 2014, Springer. p. 305-315.

[19] Luo, J., W. Wang, and H. Qi, Spatio-temporal feature extraction and representation for RGB-D human action recognition. Pattern Recognition Letters, 2014.

[20] Karg, M. and A. Kirsch, Simultaneous Plan Recognition and Monitoring (SPRAM) for Robot Assistants, in In Proceedings of Human Robot Collaboration Workshop at Robotics Science and Systems Conference 2013.

[21] Dell'Acqua, P., et al. An assistive tool for monitoring physical activities in older adults. in Serious Games and Applications for Health (SeGAH), 2013 IEEE 2nd International Conference on. 2013. IEEE.

[22] Elangovan, V., V.K. Bandaru, and A. Shirkhodaie, Team Activity Analysis and Recognition Based on Kinect Depth Map and Optical Imagery Techniques. 2012.

[23] Yang, S. and B. Li, A novel and stable human detection and behavior recognition method based on depth sensor. 3D Research, 2013. 4(2): p. 1-11.

[24] Hartley, R. and A. Zisserman, Multiple view geometry in computer vision. 2003: Cambridge university press.

[25] Chen, L., H. Wei, and J. Ferryman, A survey of human motion analysis using depth imagery. Pattern Recognition Letters, 2013. 34(15): p. 1995-2006.

[26] Kinect for Windows features. 2014; Available from: http://www.microsoft.com/en-us/kinectforwindows/discover/features.aspx.

[27] JointType Enumeration. 2013; Available from: http://msdn.microsoft.com/en-us/library/microsoft.kinect.jointtype.aspx.

[28] Karg, M. and A. Kirsch. A human morning routine dataset. in Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems. 2014. International Foundation for Autonomous Agents and Multiagent Systems.

[29] Ordónez, F.J., P. de Toledo, and A. Sanchis, Activity recognition using hybrid generative/discriminative models on home environments using binary sensors. Sensors, 2013. 13(5): p. 5460-5477.

[30] Kim, E., S. Helal, and D. Cook, Human activity recognition and pattern discovery. IEEE Pervasive Computing, 2010. 9(1).

[31] Lei, J., X. Ren, and D. Fox. Fine-grained kitchen activity recognition using rgb-d. in Proceedings of the 2012 ACM Conference on Ubiquitous Computing. 2012. ACM.

[32] Borenstein, G., Making Things See: 3D vision with Kinect, Processing, Arduino, and MakerBot. 2012: " O'Reilly Media, Inc.".

[33] Maimone, A. and H. Fuchs. Encumbrance-free telepresence system with real-time 3D capture and display using commodity depth cameras. in Mixed and Augmented Reality (ISMAR), 2011 10th IEEE International Symposium on. 2011. IEEE.

[34] Saha, S., et al., Neural Network Based Gesture Recognition for Elderly Health Care Using Kinect Sensor, in Swarm, Evolutionary, and Memetic Computing. 2013, Springer. p. 376-386.

[35] Weinland, D., R. Ronfard, and E. Boyer, A survey of vision-based methods for action representation, segmentation and recognition. Computer Vision and Image Understanding, 2011. 115(2): p. 224-241.

[36] Rojas, R., AdaBoost and the super bowl of classifiers a tutorial introduction to adaptive boosting. Freie University, Berlin, Tech. Rep, 2009.

[37] Kinect for Windows SDK 2.0. 2014; Available from: http://www.microsoft.com/en-us/download/details.aspx?id=44561