Journal of Ambient Intelligence and Smart Environments 5 (2013) 443–461
DOI 10.3233/AIS-130223
IOS Press

Analyzing the impact of different action primitives in designing high-level human activity recognition systems

Atif Manzoor [a,*], Hong-Linh Truong [b], Alberto Calatroni [c], Daniel Roggen [c], Mélanie Bouroche [a], Siobhán Clarke [a], Vinny Cahill [a], Gerhard Tröster [c] and Schahram Dustdar [b]

[a] *School of Computer Science and Statistics, Trinity College Dublin, Dublin, Ireland*
[b] *Distributed Systems Group, Vienna University of Technology, Vienna, Austria*
[c] *Wearable Computing Laboratory, Swiss Federal Institute of Technology Zürich, Zürich, Switzerland*

Abstract. Designing human activity recognition systems, an integral part of any ambient assisted living environment, is an active area of research in the ubiquitous computing, wearable sensing, and computer vision communities. Yet most of the systems ignore human body motion and arm motion action primitives to recognize high-level human activities and are limited to object usage action primitives. Consequently, there is little understanding of the significance of these action primitives on the performance of activity recognition systems. In this paper, we comparatively assess the role of the object usage action primitives, body motion action primitives, and arms motion action primitives to recognize human activities of daily living. Our experiments show that the body motion action primitives and arms motion action primitives are vital to recognize the human activities that do not involve much interaction with the objects and the environment.

Keywords: Human activity recognition, ambient assisted living, smart homes, sensing systems, machine learning

1. Introduction

Ambient assisted living environments aim to assist people by adapting to their requirements and supporting them throughout the day. Human activity is the fundamental type of context to build any application in such environments [3,12,25]. Consequently, designing dependable human activity recognition systems is an active field of research in many computer science communities. These communities use different type of sensing systems, such as ambient sensing systems [9,19,31], vision based sensing systems [1,29], and wearable sensing systems [15,37], to observe smart home environments. Learning algorithms reason with sensor data to detect action primitives or low-level human activities that span over a very short period of time, such as walk, stand, cut, release, use plate, or open fridge door [11,26]. These low-level human activities or action primitives are further used to detect high-level human activities of daily living (ADL) that span over comparatively longer periods of time, such as preparing dinner, eating sandwich, and cleaning room [7,33]. Still, one of the key challenges in building effective and reliable high-level human activity recognition systems is to identify an optimal set of action primitives that express enough information to accurately recognize such ADL.

Identifying the significance of specific action primitives in recognizing a particular ADL is important for two reasons. First, it will help designing an ambient intelligent environment by indicating where to place sensors, such as on the body, in the objects, or in the environment. Second, it also indicates which action primitives are worthwhile to invest additional effort in designing action primitive spotting algorithms. For example, "hand cutting the bread movement" could be an

*Correspinding author. E-mail: atif.manzoor@scss.tcd.ie.

important action primitive as it may give a clear indication that a subject is preparing a sandwich. Categorizing significant action primitives will also decrease the overhead in terms of sensor cost, human effort, and computing resources. But, there is little research in identifying the impact of various types of action primitives in recognizing ADL. Most work has been limited to use a specific kind of sensing system and sticks with a certain type of action primitives to recognize high-level human activities.

In this paper, we utilize human body and arms motion action primitives to recognize ADL along with commonly used object usage action primitives. In our experiments, we use different types of action primitives separately and in combinations with each other to analyze their impact. Along with defining those combinations on the basis of hand-picked action primitive types, we also use feature selection algorithms to choose meaningful action primitives. In our experiments, we use the annotations of the EU project OPPORTUNITY data set [27], which are based on the recordings of the proceedings of data collection activity as action primitives. These annotations include body movement primitives, such as walk, sit, and stand; arm movement primitives, such as reach, release, and cut; and object or environment usage primitives, such as use glass, move chair, and open fridge door.

This work is based on our previous effort to identify important action primitives to recognize high level human activities [22] and has the following major contributions.

- Uses the human body and arms motion action primitives to recognize high-level human activities
- Analyzes the performance of different classification algorithms to recognize ADL using different combination of object usage, body motion, and arm motion action primitives
- Evaluates the impact of different choices of action primitives on recognizing every individual human activity of daily living
- Provides recommendations to choose among different options of sensing systems and their placement to design and deploy an ambient assisted living environment
- Recommends to use human wearable sensors to improve the performance of human activity recognition systems and broad the range of activities predicted by these systems
- Recommends to use wearable sensors on the subject's dominant arms

The rest of the paper is organized as follows. Section 2 provides an overview of existing work to recognize human activities of daily living. Section 3 briefly describes the EU project OPPORTUNITY data set used in this work. Section 4 presents an overview of our methodology which includes our data processing and feature extraction technique, action primitive sets, algorithms used in this work, and our evaluation criteria and strategy to measure the performance of classifiers to correctly predict human activities. Section 5 presents our experiments and their results. Section 6 details the discussion of our results and provides recommendations. Finally, Section 7 concludes the paper.

2. Related work

Human activity is an important type of context to develop any ambient assisted living application [21]. Consequently considerable research effort has been undertaken to recognize ADL. This research uses different approaches and sets of sensors to achieve those objectives. The wearable sensing community has used body-worn accelerometers to detect low-level human activities, such as lying, sitting, standing, walking, running, and cycling [15,37]. The ubiquitous computing researchers embedded the environment with different kind of sensors, such as reed switches and RFID tags, to collect information about object usage in the environment and recognize high-level human activities such as eating, drinking, cooking, grooming, and leaving home [8,33]. The computer vision community used video camera sensors to capture and predict current human activity, ranging from low-level body movements or action primitives to high-level human activities of daily living [1,18,29]. These works also used a wide range of algorithms and evaluation criteria to present the effectiveness of their approaches. Here, we discuss and compare this research with our work, with a summary presented in Table 1.

D.J. Cook [8] used naive Bayes, hidden Markov model, and conditional random field to learn setting-generalized human activity recognition models for smart homes. They evaluated their approach using eleven separate data sets from CASAS[1] Smart Home project [10]. They have recognized ADL related to

[1] http://ailab.wsu.edu/casas/

Table 1
Research works performing human activity recognition

Work references	Action primitives				Sensor systems								Sensor placement				Evaluation metrics						
	Body motions	Arm motions	Object usage	Others	Acceleration	Power	Reed switches	RFID	Positioning	Audio	Visual	Others	Torso	Limbs	Environment	Objects	Accuracy	True positive rate	False positive rate	Precision	F-measure	ROC area	Confusion matrix
This work	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
D. J. Cook 2012 [8]	×	×	✓	×	✓	✓	✓	✓	×	×	×	✓	×	×	✓	✓	✓	×	×	×	×	×	✓
Rashidi, et al. 2012 [25]	×	×	✓	×	✓	✓	✓	✓	×	×	×	✓	×	×	✓	✓	✓	×	×	×	×	×	✓
Wang, et al. 2011 [33]	✓	✓	✓	✓	✓	✓	✓	×	✓	✓	×	✓	✓	✓	✓	✓	✓	×	×	×	×	×	✓
Gu, et al. 2011 [13]	✓	✓	✓	✓	✓	✓	✓	×	✓	✓	×	✓	✓	✓	✓	✓	✓	×	×	×	×	×	✓
Lee, et al. 2011 [17]	×	×	×	×	✓	×	×	×	×	×	×	×	✓	✓	×	×	✓	×	×	×	×	✓	×
van Kastern, et al. 2010 [30]	×	×	✓	×	×	×	×	×	×	×	×	×	×	×	✓	✓	✓	×	×	×	×	×	✓
Lepri, et al. 2010 [18]	✓	×	×	✓	×	×	×	×	×	×	✓	×	×	×	✓	×	✓	×	×	×	✓	×	✓
Maekawa, et al. 2010 [20]	×	×	✓	✓	✓	×	×	×	×	✓	✓	✓	×	✓	×	×	✓	×	×	×	×	×	✓
Ye, et al. 2010 [36]	×	✓	✓	✓	✓	✓	✓	✓	×	✓	✓	✓	×	✓	✓	✓	×	✓	×	✓	✓	×	×
Mckeever, et al. 2009 [23]	×	×	✓	×	✓	×	✓	✓	×	×	×	×	×	×	✓	✓	✓	×	×	×	×	×	✓
Choudhury, et al. 2008 [6]	×	×	✓	×	✓	×	×	×	✓	×	✓	×	✓	✓	×	×	✓	✓	✓	✓	×	×	×
Logan, et al. 2007 [19]	×	✓	✓	✓	✓	✓	✓	✓	×	✓	✓	✓	×	✓	✓	✓	×	✓	✓	×	×	✓	✓
Wu, et al. 2007 [34]	×	×	✓	×	×	×	×	✓	×	×	✓	×	×	✓	×	✓	×	×	×	×	×	×	×
Tapia, et al. 2004 [28]	×	×	✓	×	×	×	✓	×	×	×	×	✓	×	×	✓	✓	✓	×	×	×	×	×	✓

personal hygiene, grooming, taking meals, and eating. Rashidi et al. [25] introduced an unsupervised method to recognize and track human activities in smart environments, to monitor their health conditions and provide them with medical assistance. They also used the CASAS Smart Home project data sets in their work. However, CASAS Smart Home project data sets are limited to only environment and object usage action primitives and analyzing the impact of different kinds of action primitives on high-level human activity recognition was not the main target of these research efforts.

Wang et al. [33] investigate the problem of recognizing multiuser activities using wearable sensors in a smart home setting. They developed a multi-modal wearable sensing platform to collect sensor data about multiple users. They packaged accelerometers, audio, and RFID sensors in their sensing platform. Users wore this sensing platform on their limbs and torso. Gu et al. [13] propose a novel pattern mining approach to recognize sequential, interleaved, and concurrent activities in a unified framework. Although, both of these works used almost the same sensing modalities as our work, they did not perform any analysis of the performance of the sensing modalities in detecting human activities. They also did not investigate the optimal position of sensing systems on the human body and in the environment.

Lee et al. [17] developed a portable personal life log system. They used a 3-axial accelerometer to detect human activities, such as lying, standing, and walking. Tapia et al. [28] used environmental state change sensors that had been installed on doors, windows, cabinets, and drawers etc. Kasteren et al. [30] also stress the need of a non-intrusive system to recognize human activities to design an automatic health monitoring system for elderly people. They collected their data set using a wireless sensor network consisting of reed switches, pressure mats, mercury contacts, passive infrared, float sensors, and temperature sensors. The main goal of their system is to show the effectiveness of generative and discriminative models for activity recognition in a real world setting. They compared hidden Markov model and conditional random field for that purpose. They did compare the performance of the aforementioned models using different type of sensor data.

Mckeever et al. [23] used a publicly available data set [31] to recognize human activities of daily living. In their approach, they modeled domain knowledge and transfered the evidence from the sensor to situation level to minimize the dependence on the training data sets. Ye et al. [35,36] used domain knowledge and machine learning algorithm to develop situation lattices to recognize ADL. First, they used the domain knowledge to map the sensor data to low-level situation semantics or action primitives. Later, they used machine learning algorithms to map low-level situation semantics or action primitives to high-level situation semantics or human activities of daily living. They used the publicly available data sets, place lab [16] and Kasteren et al. [31], to evaluate their approach. Comparing with decision tree and naive Bayes, their algorithms showed a varying degree of performance with considerably low precision to recognize ADL using a sensor-rich data set. They identified that lack of domain knowledge leading to imprecise situation or activity definition may be one of the reason of low precision. Our work that identify the impact of different action primitives in designing high-level human activity recognition systems may provide enough domain knowledge to construct such situation lattices.

Lepri et al. [18] recognized ongoing activities by using visual sensors. They stressed that the current human body posture may be useful to distinguish between different human activities that involve the same kind of objects, such as eating and preparing meals. They equipped the living room and kitchen of a flat with web cameras where different subjects performed activities of daily living. They processed the video streams to get the primitives about the location and posture of the subject to recognize high-level activities. But these primitives did not prove to be enough to recognize activities like eating, drinking, and cleaning. Most aforementioned approaches used sensor systems that provide primitives about the environmental state change. Lepri et al. [18] used human body posture primitives with their location to detect ADL. Choudhury et al. [6] stressed that human activity recognition using body-worn sensors is crucial to build many health-care applications, such as fitness monitoring and elder-care support. Their work is comparatively close to our work as they packaged multiple sensors, such as microphones, accelerometer, and light sensors, in a single small device and tried to understand the usefulness of those sensors in recognizing different human activities. However, compared to our work, their sensor set is quite limited and they have not included environment and object usage sensors in their study. Their human activity set is also limited to low-level human activities, such as walking, sitting, and standing.

Maekawa et al. [20], Logan et al. [19], and Wu et al. [34] compared the performance of different sensing systems to recognize human activities of daily living and are close to our work. Maekawa et al. [20] recognized activities of daily living with their custom-designed device embedded with a camera, microphone, accelerometer, and a digital compass. The subjects wore the device on their wrist and performed different activities of daily living. They compared the performance of each sensor in data set by including and excluding each type of sensor. Their results assert that a video camera is the best sensor to use in a human activity recognition system. However, they also admitted that there may be many privacy issues concerning the use of cameras and microphones in the home setting. RFID sensors and accelerometers are also cheap to deploy. In addition, they did not use enough accelerometers to recognize human activity. In our work, accelerometer sensors are used to recognize body motion and arm motion action primitives, which proved vital to recognize human activities. Compared to their work, we put more emphasis on comparing the action primitives and used a wide range of action primitives, as evident from Table 1.

Logan et al. [19] used the place lab, an instrumented home environment [16], to collect their data set. They used environment built-in sensors, object motion sensors, and RFID tags. Two 3-axis accelerometer sensors were also worn by a subject on his limbs to show his motion. In their experiments, they compared the performance of RFID sensors, motion sensors, and built-in sensors to recognize ADL. However, all these different kinds of sensors were used to extract object-usage action primitives. Consequently, they found that the activities that do not involve any object usage are hard to estimate, such as reading. In contrast to their work, we emphasize investigating the impact of action primitives to recognize ADL. We also use a comprehensive set of action primitives that include object usage action primitives, body motion action primitives and arm motion action primitives.We analyze the impact of these action primitives on each individual activity separately. Body motion action primitives and arm motion action primitives proved vital in recognizing activities that do not use many objects in the environment.

Wu et al. [34] recognized human activity using object usage action primitives. They used RFID sensors and video cameras to predict object usage action primitives. They compared the performance of RFID sensors and video cameras to detect an object usage and subsequently high level human activities of daily living, such as boil water, make popcorn, and make tea. Compared to our work they target the recognition of kitchen activities that generate a high number of object usage action primitives.We also use body motion and arms motion action primitives that prove vital in recognizing activities that do not involve much interaction with objects in the environment.

Most of the aforementioned research efforts were concerned with either comparing the effectiveness of existing classification algorithms or designing new more effective classification algorithms to predict ADL. They used different sensing systems for that purpose. Some of them also compared the effectiveness of the sensing system to recognize ADL. However, little attention is paid to the comparison of the action primitives. Compared to these works, we analyze the impact of different choices of action primitives on recognizing ADL. We also discuss our results in the context of the choice and placement of sensing systems for the design of an effective activity recognition system in an ambient assisted living environment.

3. Data description

In our experiments, we use the smart home data set gathered in the EU project OPPORTUNITY[2] for the machine recognition of human activities of daily living. Here, we briefly describe the features of the data set that are related to our work. Interested readers may see the detailed description of the data set in [27].

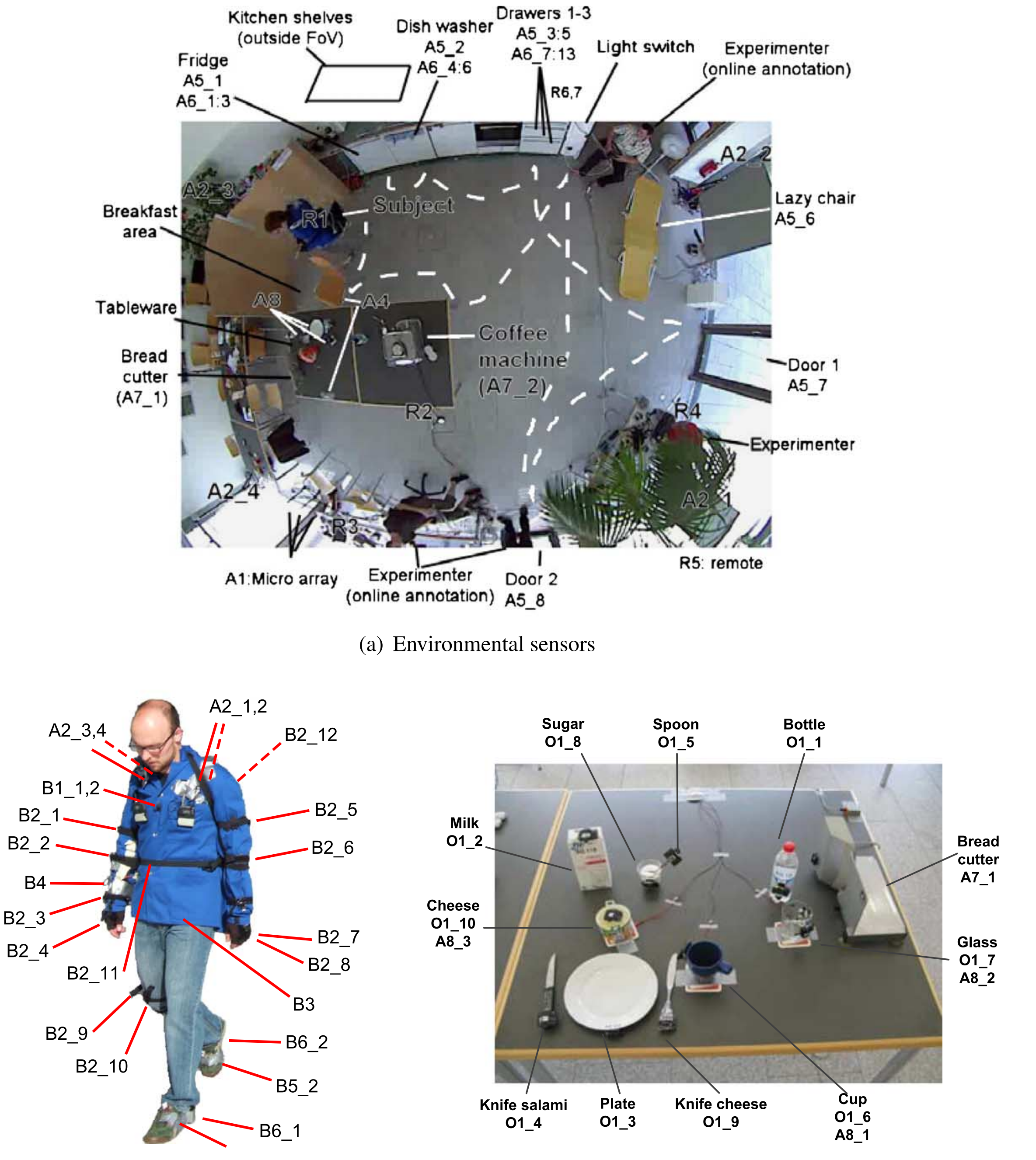
The data set is collected in a sensor-rich environment: a room simulating a studio flat with kitchen, deckchair, and outdoor access where subjects performed daily morning activities. Fifteen networked sensor systems with seventy two sensors of ten modalities were deployed in the environment, embedded in the objects, and worn on the subject body. Table 2 shows the description and location of these sensor systems. Figure 1 shows the detail of the deployment of sensing systems. Figure 1(a) shows the sensors that were deployed at different places in the room. Figure 1(b) shows the location of the sensors that were worn on the subject body. Figure 1(c) shows the sensors that were embedded in different objects. Twelve subjects executed activities of daily living in this environment, yielding an average of 2 hours of effective data per subject, for a total of twenty five hours of sensor data. According to estimations over 11000 interactions with objects and over 17000 interactions with environment have been recorded. This makes the data set highly rich in gesture primitives and largest for the purpose of multi-modal activity recognition.

Table 2
Locations and observations of deployed sensor systems

ID	Sensor system	Location and observation
B1	Commercial wireless microphone	Chest and dominant wrist. Senses user activity
B2	Custom wireless Bluetooth acceleration sensors	12 locations on the body. Senses limb movement
B3	Custom motion jacket	Jacket including 5 commercial RS485 networked sensors. Sense inertial measurement units
B4	Custom magnetic relative positioning sensor	Emitter on shoulder, receiver on dominant wrist. Senses distance of hand to body
B5	Commercial Inertia Cube3 inertial sensor system	One per foot, on the shoe toe box. Senses modes of locomotion
B6	Commercial Sun SPOT acceleration sensors	One per foot, right below the outer ankle. Senses modes of locomotion
O1	Custom wireless Bluetooth acceleration and rate of turn sensors	On 12 objects used in the scenario. Senses object use
A1	Commercial wired microphone array	4 at one room side. Senses ambient sound
A2	Commercial Ubisense localization system	Corners of the room. Senses user location
A3	Axis network cameras	3 locations, for localization, documentation and visual annotation
A4	XSense inertial sensor	On the table and chair. Senses vibration and use
A5	USB networked acceleration sensors	USB networked acceleration sensors. Sense usage
A6	Reed switches	13, on doors, drawers, shelves. Sense usage, provides ground truth
A7	Custom power sensors	Connected to coffee machine and bread cutter. Senses usage
A8	Custom pressure sensors	3 on the table, user placed plates and cups on them. Senses usage

The data set is annotated at two levels of abstraction. At a high level of abstraction, annotators label the data set with ADL. Table 3 shows the short description of these activities and their duration for a single run of the data set. A subject starts the ADL run with Idle activity. During this activity the subject is lying on the deckchair. During the *Relaxing* activity the sub-

[2]http://www.opportunity-project.eu/



(a) Environmental sensors

(b) Body sensors

(c) Object sensors

Fig. 1. Detail of the sensing systems used to collect data from the environment. Figure 1(a) shows sensing systems deployed in the environment. Figure 1(b) shows sensing system worn on the human body and Fig. 1(c) shows sensing systems embedded in the objects. Table 2 provides further detail about the sensing systems.

ject gets up, goes outside the room, and has a walk around the building. During the *Early Morning* activity the subject moves around the room and casually checks the different objects. Later, the subject prepares coffee with milk and sugar using a coffee machine and drinks coffee during the *Coffee Time* activity. The subject uses bread, cheese, salami, bread cutter, knives, and plates to prepare and later eat the sandwich dur-

Table 3
Activities and their duration during a single run

Activities	Description	Duration (seconds)
Idle	Lie on the deckchair in the room	583
Relaxing	Go outside and have a walk	157
Early Morning	Move around in the room and casually check the objects	276
Coffee Time	Prepare coffee with milk and sugar using coffee machine and drink it	129
Sandwich Time	Prepare sandwich with bread, cheese, and salami using bread cutter, various knives, and plates and eat it	375
Clean Up	Put objects used to original place or dish washer and cleanup the table	183

Table 4
Brief description and values of action primitive categories

Action primitive category	Description	Primitive Values
Body motions	Basic human movements	walk, run, stand, lie, sit, stairs up, stairs down
Left arm motions	Left arm movements	reach, move, release, lock, unlock, open, close, stir, sip, bite, clean, cut, spread
Right arm motions	Right arm movements	
Left arm object usage	Left hand interaction with objects	fridge, dishwasher, drawer1 (top), drawer2 (middle), drawer3 (lower), door1, door2, switch, table, cup, chair, glass, spoon, sugar, knife salami, knife cheese, salami, bottle, plate, cheese, bread, milk, lazy chair
Right arm object usage	Right hand interaction with objects	

ing the *Sandwich Time* activity. At the end the subject puts all the used objects to their original positions or the dish washer and cleans the table in the *Clean Up* activity.

At a low level of abstraction, annotators label the data set with action primitives. Table 4 shows the detail of action primitives used to label the data set. Body motion action primitives include action primitives that present information about the subject body movements, such as walking, sitting, and lying. Arms motion action primitives include action primitives that represent information about the actions performed by human limbs on different objects, such as cut, spread, and release. Object usage action primitives represent information about the usage of different objects in the environment, e.g., plates, cups, and knives.

In our experiments, we use the data set labels at the two aforementioned levels of abstraction. Classifiers use the low level labels of action primitives as the feature set to classify high level labels of ADL. We use the data of fifteen runs of the ADL. Three different users performed five runs each of the fifteen runs. Each run last for approximately thirty minutes. We use total data of about four hundred and fifty minutes.

4. Methodology

Figure 2 shows our system to recognize human activities of daily living. The system collects data from the sensors embedded in the environment and objects and worn by the human. Machine learning algorithms are used to extract initially action primitives from sensor data and later human activities from action primitives. The system may also fuse the output of different learning algorithms to improve the prediction accuracy. However, the scope of this work is limited extracting the high-level human activity of daily living from the action primitives as shown by the dotted box in Fig. 2. In this section, we describe the data processing and feature extraction techniques, action primitive sets, evaluation criteria and strategy, and algorithms that we use in the paper.

4.1. Data processing and feature extraction

We use the sliding window technique to extract feature histograms from action primitives temporal data. The first step in using this technique is to define a window size, W, of a certain time length and process data collected in that window. For example, as shown in Fig. 3, we define a window size, W, of five seconds for the available action primitive temporal data. We count the frequency of a specific action primitive label during that window and construct the feature histogram as shown in Fig. 3. Feature histograms indicate the number of times an action primitive is labeled in that interval of time. For example, if the data indicates that the subject used a plate seven times in that time period, the feature "plate used" will have the value seven for that time period. We pass the feature vectors of different action primitives to the classification algorithms to classify the subject ADL during this interval of time. The second step is to mention the window jump size. For example, we define a jump size, JS, of one sec as shown in Fig. 3. For every subsequent sampling, we move the window ahead with the length of the jump

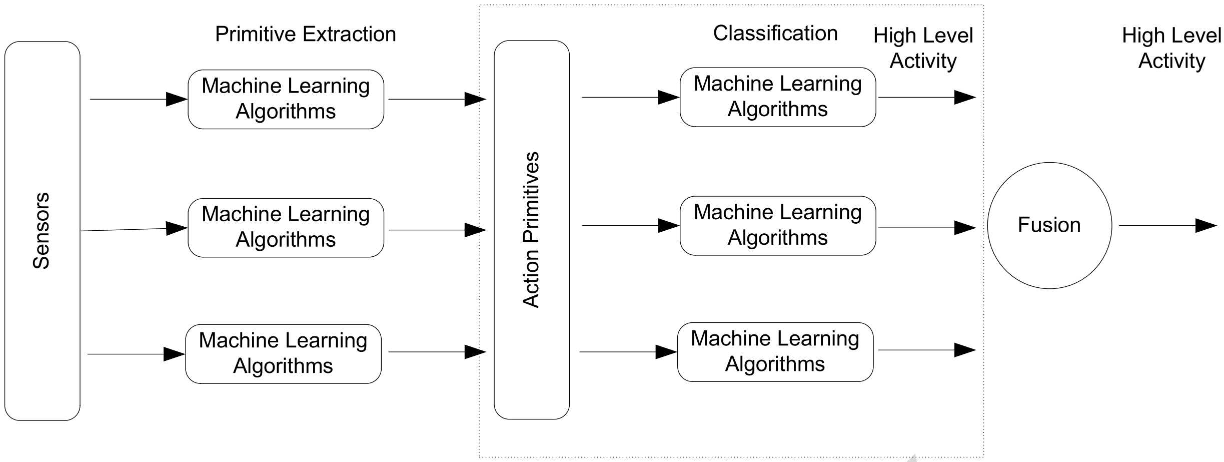


Fig. 2. The system to recognize human activities of daily living collect data from sensing systems deployed in the environment and use machine learning algorithms to extract action primitives from sensor data and human activities of daily living from action primitives at the later stage. The system may also fuse the output from different learning algorithms to improve the accuracy of the system.

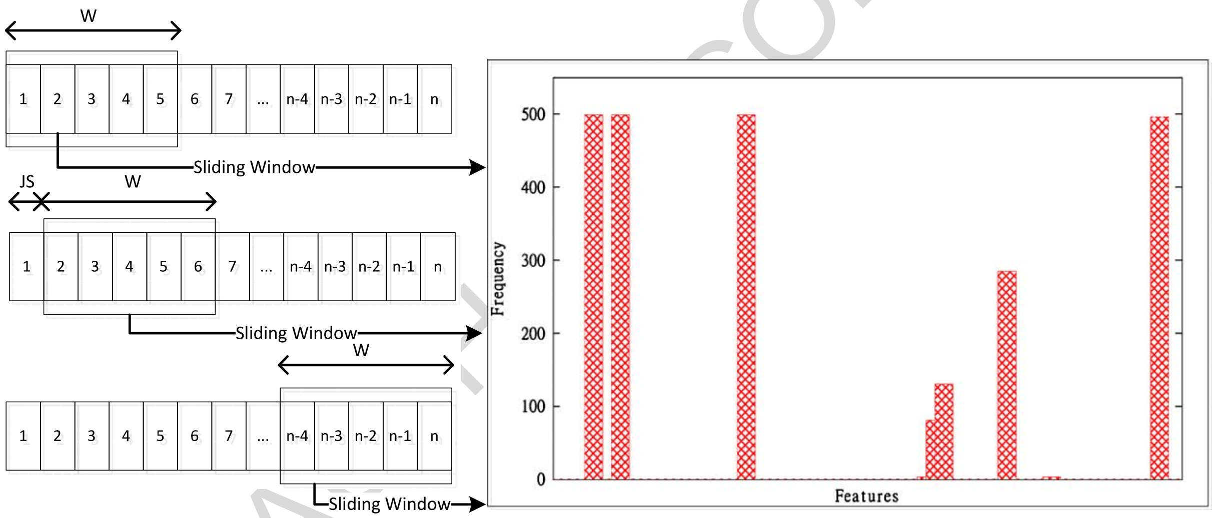


Fig. 3. Sliding window technique to extract features from raw data. First, we define a window size, W, of five seconds and jump size, JS, of one second for a temporal signal of n seconds. Second, we count the frequency of sensors fired during that window and construct the feature histogram. For every subsequent sampling we move the sliding window ahead with jump size of one second.

size. For every next sample, data is taken from a next window of length five seconds starting one second after the start of the previous window until the sliding window reaches the end of the signal of length n seconds.

In our experiments, we use jump size, JS, of one second and window sizes, W, of one, five, ten, thirty, and sixty seconds. First, we find out the optimal window size for our data. We persist with that optimal window size for the remaining part of our experiments.

4.2. Action primitive sets

A person performs different activities in different parts of a house. He/she may be busy in the kitchen while preparing breakfast or he/she may be relaxing in the lounge. The different nature of these activities implies that these activities are composed of different primitives. The composition of these activities requires us to look at each activity individually which will not only give us the opportunity to observe which type of sensors should be used to recognize which activ-

Table 5
Detail of different sets of action primitives, categories of action primitives used in these sets, and sensing systems used to extract these action primitive categories

Set	Categories	Sensing system[3]
$S1$	body motions, arm motions, object usage	A4, A5, A6, A7, B1, B2, B3, B4, B5, B6, O1
$S2$	body motions, arm motions	B1, B2, B3, B4, B5, B6
$S3$	body motions	B1, B3, B4, B5, B6
$S4$	arm motions	B2, B4
$S5$	right arm motions	B2, B4
$S6$	left arm motions	B2, B4
$S7$	object usage	O1, A4, A5, A6, A7
$S8$	Feature Selection	A4, A5, A6, A7, B1, B2, B3, B4, B5, B6, O1

ity but will also indicate which type of sensors should be used in which part of house. Considering these requirements, we also divide action primitives in different combinations. These different combinations provide us the opportunity to study which primitives are more suitable to predict a specific type of human activity of daily living.

Table 5 shows the action primitive sets, the categories of action primitives that are used in those sets, and the sensing systems used to extract those action primitives. In the set *S1* we use all the action primitive categories described in Table 4. *S2* consists only body motion action primitives and arms motion action primitives. In *S3* we further excluded arm motion action primitives and it only consists of body motion action primitives. In *S4* we use both right arm motion and left arm motion action primitives while *S5* and *S6* are limited to only right arm motion and left arm motion respectively. In *S7* we only use the object usage action primitives. In *S8* we use correlation-based feature subset selection [14] as the evaluation criteria and linear forward selection as the search algorithm. The detail of the values that can be assigned to these action primitives is shown in Table 4.

4.3. Algorithms

We use the WEKA[4] implementations of decision tree [24], Bayes net [5], and k-nearest neighbors [2]. These classification algorithms are commonly used in previous reported works on human activity recognition [4,19,26]. We have chosen them as the purpose of our research is to assess whether we can find a set of action primitives yielding high performance to predict ADL regardless of the classification strategy.

4.4. Evaluation strategy and criteria

Mostly activity recognition algorithms in smart homes are evaluated considering the accuracy of the algorithms that represent the ratio of correctly predicted activities against the total number of activities. However, the accuracy of an algorithm is not enough to evaluate the real performance of a classifier. For example, a classifier is used to classify a set of instances that has nine positive instances and one negative instances. If the classifier classifies all the instances as the positive instance, it will have ninety percent accuracy and hundred percent false positive rate. So accuracy alone does not show the true picture of the performance of the classifier. In this paper, we use different evaluation metrics, such as recall, false alarm rate, precision, f-measure, confidence intervals, and kappa statistical measure [32] to evaluate the performance of classification algorithms. We use confusion matrix as a tool to calculate these metrics.

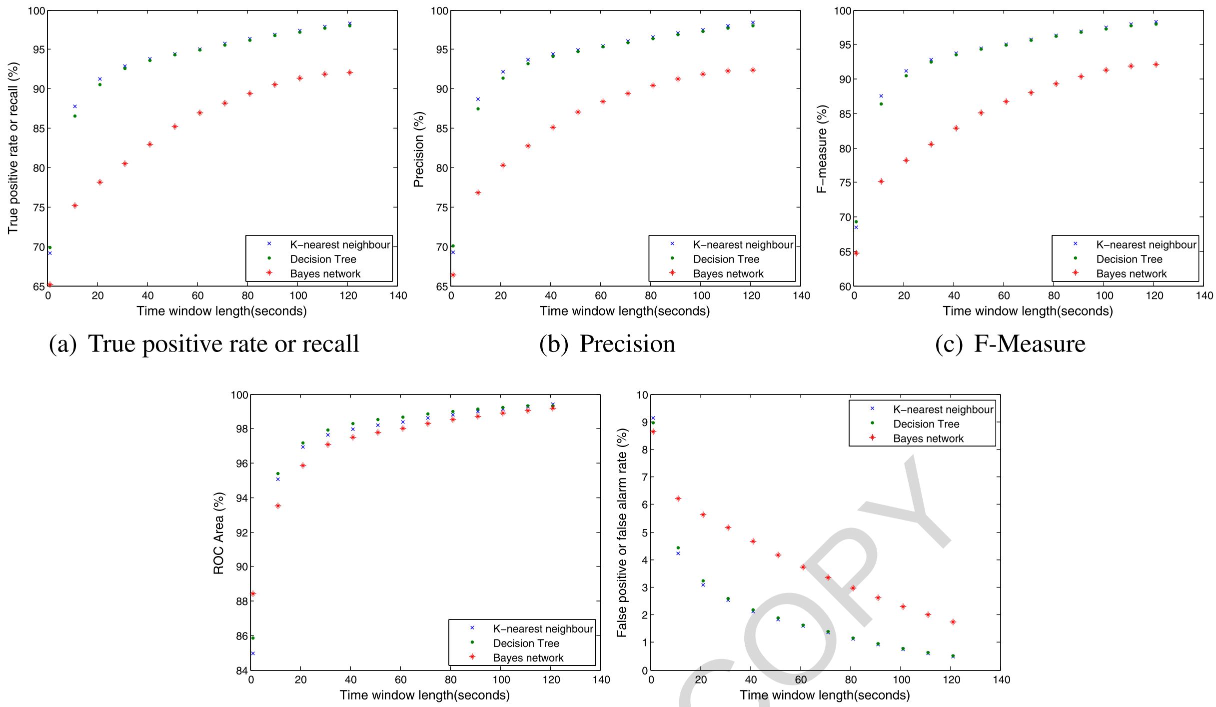
5. Experiments

In our experiments, classifiers take the action primitives, listed in Table 4, as input to the system and predict the ADL, listed in Table 3, as output to the system. We first assess the suitable window size to extract feature histograms over the temporal signal (Section 5.1). We use the assessed suitable window size in the remaining experiments. Afterwards, we analyze the impact of using different action primitives sets described in Table 5 on recognizing each individual activity listed in Table 3 (Section 5.2). Finally, we analyze overall impact of action primitive sets on activity recognition (Section 5.3).

5.1. Impact of a change in sliding window size on acitivity recognition

In this experiment, we analyze the effect of the sliding window length to extract feature histograms of the action primitives over the prediction of ADL. We use the time window lengths in the range of one second to two minutes. Classifiers take feature histograms ex-

[3]Sensing systems are described in Table 2.
[4]http://www.cs.waikato.ac.nz/ml/weka/index.html



(a) True positive rate or recall (b) Precision (c) F-Measure

(d) ROC Area (e) False positive or false alarm rate

Fig. 4. Different evaluation metrics for using k-nearest neighbour, decision tree, and Bayes net classifier with different time window lenghts.

tracted over these window lengths to predict ADL. Figure 4 shows the evaluation metrics for the classification algorithm k-nearest neighbor, decision tree, and Bayes network to predict all the activities collectively. We observe that the value of the true positive rate or recall (Fig. 4(a)) increases and the value of false positive rate or false alarm rate (Fig. 4(e)) decreases with the increase in the size of window length. Similarly precision (Fig. 4(b)), f-measure (Fig. 4(c)), and roc area (Fig. 4(d)) also have better values with the increase in the window size. We also observe that the evaluation metrics show the same pattern with all three classification algorithms.

Figure 4 also shows that the slope of the curves plotting the evaluation metrics become flatter with the increase in the sliding window length and the improvement in the performance of the classification algorithms to correctly classify the human activity is almost negligible after a certain extent. So it is not very useful to increase the sliding window length beyond a certain size. Large sliding window size may also prevent the algorithms to recognize the activities that last for a short period of time, e.g., if we have a sliding window size of two minutes, it will be difficult to recognize an activity that last for only sixty seconds. Smart home applications will also wait for the sliding window length time before a certain human activity in a smart home can be recognized. Such delay can affect the functionality of those applications. Consequently, in spite of increasing the sliding window size beyond certain limit, it is recommended to look at other options to improve the performance of activity recognition algorithms. Considering these observations, we decide to use a window length of sixty seconds for the subsequent experiments.

5.2. *Impact of different action primitive sets on recognizing each activity*

In this experiment, we examine the impact of action primitive sets on each activity separately. We use the sliding window length of sixty seconds. Figure 5 shows the percentage value of true positive rate with confidence intervals for the prediction of different activities using Bayes net classifier. Similarly Figs 6 and 7 show the percentage value of true positive rate with confidence interval for the prediction of different activities using a decision tree and a k-nearest neighbor classification algorithms. Table 6 shows the detailed performance of all classifiers with each ac-

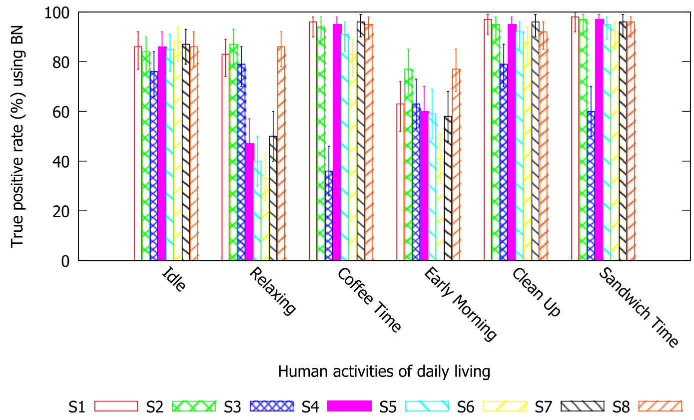


Fig. 5. True positive rate (%) with confidence interval to recognize different activities using Bayes net classifier with different primitive sets.

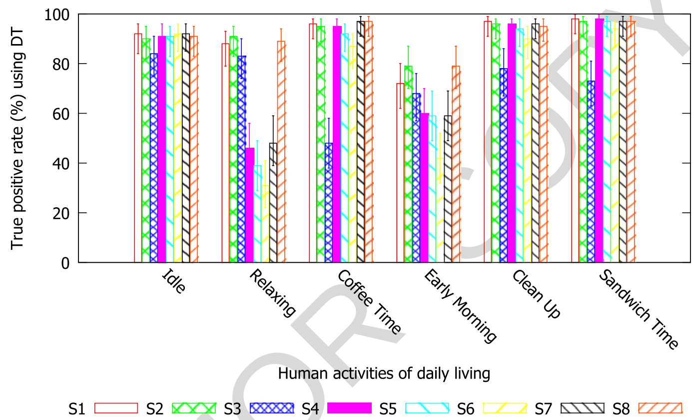


Fig. 6. True positive rate (%) with confidence interval to recognize different activities using decision tree classifier with different primitive sets.

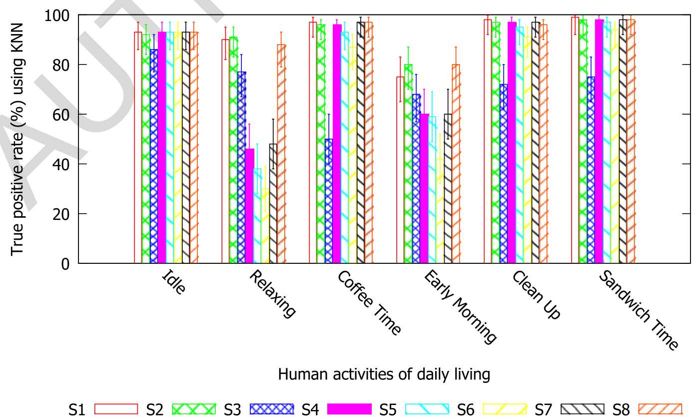


Fig. 7. True positive rate (%) with confidence interval to recognize different activities using k-nearest neighbors classifier with different primitive sets.

tion primitive set to predict these activities. The performance measure from those confusion matrices are used to calculate the percentage value of true positive rate and confidence intervals. We used the confidence level of 95% to estimate confidence interval.

Table 6

Confusion matrices for different classifiers and primitive sets

(a) BN with primitive set 1

	Predicted Activities (%)					
True Activities	*Idle*	*Relaxing*	*Coffee time*	*Early morning*	*Clean up*	*Sandwich time*
Idle	86	7	2	1	3	1
Relaxing	17	83	0	0	0	0
Coffee time	3	0	96	1	0	0
Early morning	37	0	0	63	0	0
Clean up	0	0	0	0	97	3
Sandwich time	1	0	1	0	0	98

(b) DT with primitive set 1

	Predicted Activities (%)					
True Activities	*Idle*	*Relaxing*	*Coffee time*	*Early morning*	*Clean up*	*Sandwich time*
Idle	92	4	1	1	1	1
Relaxing	11	88	0	1	0	0
Coffee time	2	0	96	1	0	1
Early morning	27	1	0	72	0	0
Clean up	1	1	0	0	97	1
Sandwich time	1	0	1	0	0	98

(c) KNN with primitive set 1

	Predicted Activities (%)					
True Activities	*Idle*	*Relaxing*	*Coffee time*	*Early morning*	*Clean up*	*Sandwich time*
Idle	93	3	1	1	1	1
Relaxing	10	90	0	0	0	0
Coffee time	2	0	97	1	0	0
Early morning	24	0	1	75	0	0
Clean up	1	0	0	0	98	1
Sandwich time	1	0	0	0	0	99

(d) BN with primitive set 2

	Predicted Activities (%)					
True Activities	*Idle*	*Relaxing*	*Coffee time*	*Early morning*	*Clean up*	*Sandwich time*
Idle	84	7	3	1	3	2
Relaxing	12	87	0	1	0	0
Coffee time	3	0	94	2	0	1
Early morning	21	1	1	77	0	0
Clean up	0	1	1	0	95	3
Sandwich time	1	0	1	0	1	97

(e) DT with primitive set 2

	Predicted Activities (%)					
True Activities	*Idle*	*Relaxing*	*Coffee time*	*Early morning*	*Clean up*	*Sandwich time*
Idle	90	4	2	1	2	1
Relaxing	9	91	0	0	0	0
Coffee time	2	0	95	2	0	1
Early morning	19	1	1	79	0	0
Clean up	1	1	0	0	96	2
Sandwich time	1	0	1	0	1	97

(f) KNN with primitive set 2

	Predicted Activities (%)					
True Activities	*Idle*	*Relaxing*	*Coffee time*	*Early morning*	*Clean up*	*Sandwich time*
Idle	92	4	1	1	1	1
Relaxing	9	91	0	0	0	0
Coffee time	2	0	96	1	0	1
Early morning	19	0	1	80	0	0
Clean up	1	0	0	0	97	2
Sandwich time	1	0	1	0	0	98

(g) BN with primitive set 3

	Predicted Activities (%)					
True Activities	*Idle*	*Relaxing*	*Coffee time*	*Early morning*	*Clean up*	*Sandwich time*
Idle	76	4	4	6	8	2
Relaxing	8	79	2	3	4	4
Coffee time	6	2	36	20	26	10
Early morning	19	2	3	63	11	2
Clean up	1	1	7	4	79	8
Sandwich time	1	1	7	2	29	60

(h) DT with primitive set 3

	Predicted Activities (%)					
True Activities	*Idle*	*Relaxing*	*Coffee time*	*Early morning*	*Clean up*	*Sandwich time*
Idle	84	3	3	4	5	1
Relaxing	7	83	1	2	3	4
Coffee time	8	0	48	15	18	11
Early morning	19	1	4	67	7	2
Clean up	2	1	6	3	78	10
Sandwich time	1	1	6	2	17	73

(i) KNN with primitive set 3

	Predicted Activities (%)					
True Activities	*Idle*	*Relaxing*	*Coffee time*	*Early morning*	*Clean up*	*Sandwich time*
Idle	86	3	2	3	4	2
Relaxing	9	77	2	4	2	6
Coffee time	8	1	50	14	15	12
Early morning	20	1	4	68	5	2
Clean up	3	1	7	3	72	14
Sandwich time	2	1	5	2	15	75

(j) BN with primitive set 4

	Predicted Activities (%)					
True Activities	*Idle*	*Relaxing*	*Coffee time*	*Early morning*	*Clean up*	*Sandwich time*
Idle	86	6	2	1	3	2
Relaxing	52	47	0	1	0	0
Coffee time	3	0	95	2	0	0
Early morning	39	0	1	60	0	0
Clean up	1	1	0	0	95	3
Sandwich time	1	0	1	0	1	97

(k) DT with primitive set 4

	Predicted Activities (%)					
True Activities	*Idle*	*Relaxing*	*Coffee time*	*Early morning*	*Clean up*	*Sandwich time*
Idle	91	3	2	1	2	1
Relaxing	53	46	0	1	0	0
Coffee time	3	0	95	2	0	0
Early morning	38	1	1	60	0	0
Clean up	1	1	0	0	96	2
Sandwich time	1	0	1	0	0	98

(l) KNN with primitive set 4

	Predicted Activities (%)					
True Activities	*Idle*	*Relaxing*	*Coffee time*	*Early morning*	*Clean up*	*Sandwich time*
Idle	93	3	1	1	1	1
Relaxing	53	46	0	1	0	0
Coffee time	2	0	96	2	0	0
Early morning	38	1	1	60	0	0
Clean up	1	0	1	0	97	1
Sandwich time	1	0	0	0	1	98

(m) BN with primitive set 5

	Predicted Activities (%)					
True Activities	*Idle*	*Relaxing*	*Coffee time*	*Early morning*	*Clean up*	*Sandwich time*
Idle	85	5	3	2	3	2
Relaxing	59	40	0	1	0	0
Coffee time	3	0	91	2	1	3
Early morning	40	0	1	59	0	0
Clean up	0	1	1	0	92	6
Sandwich time	1	0	2	0	2	95

(n) DT with primitive set 5

	Predicted Activities (%)					
True Activities	*Idle*	*Relaxing*	*Coffee time*	*Early morning*	*Clean up*	*Sandwich time*
Idle	91	3	1	1	2	2
Relaxing	60	39	0	1	0	0
Coffee time	3	0	92	2	1	2
Early morning	40	0	1	59	0	0
Clean up	1	0	1	0	94	4
Sandwich time	1	0	2	0	0	97

(o) KNN with primitive set 5

	Predicted Activities (%)					
True Activities	*Idle*	*Relaxing*	*Coffee time*	*Early morning*	*Clean up*	*Sandwich time*
Idle	93	3	1	1	1	1
Relaxing	61	38	0	1	0	0
Coffee time	3	0	93	2	1	1
Early morning	40	0	1	59	0	0
Clean up	1	0	1	0	95	3
Sandwich time	1	0	1	0	1	97

Table 6

(*Continued*)

(p) BN with primitive set 6

	Predicted Activities (%)					
True Activities	*Idle*	*Relaxing*	*Coffee time*	*Early morning*	*Clean up*	*Sandwich time*
Idle	88	4	2	1	3	2
Relaxing	64	33	0	1	2	0
Coffee time	9	1	83	2	1	4
Early morning	55	1	1	41	1	1
Clean up	3	1	1	0	88	7
Sandwich time	3	0	2	1	2	92

(q) DT with primitive set 6

	Predicted Activities (%)					
True Activities	*Idle*	*Relaxing*	*Coffee time*	*Early morning*	*Clean up*	*Sandwich time*
Idle	92	3	2	1	1	1
Relaxing	67	31	0	1	1	0
Coffee time	8	1	87	1	1	2
Early morning	55	1	1	42	0	1
Clean up	2	0	1	1	90	6
Sandwich time	2	0	1	1	2	94

(r) KNN with primitive set 6

	Predicted Activities (%)					
True Activities	*Idle*	*Relaxing*	*Coffee time*	*Early morning*	*Clean up*	*Sandwich time*
Idle	93	2	2	1	1	1
Relaxing	68	30	0	1	1	0
Coffee time	8	1	87	1	1	2
Early morning	55	1	1	42	0	1
Clean up	2	0	1	1	91	5
Sandwich time	3	0	1	1	1	94

(s) BN with primitive set 7

	Predicted Activities (%)					
True Activities	*Idle*	*Relaxing*	*Coffee time*	*Early morning*	*Clean up*	*Sandwich time*
Idle	87	7	2	1	2	1
Relaxing	49	50	0	1	0	0
Coffee time	4	0	96	0	0	0
Early morning	40	1	1	58	0	0
Clean up	1	1	0	0	96	2
Sandwich time	2	0	1	0	1	96

(t) DT with primitive set 7

	Predicted Activities (%)					
True Activities	*Idle*	*Relaxing*	*Coffee time*	*Early morning*	*Clean up*	*Sandwich time*
Idle	92	4	1	1	1	1
Relaxing	50	49	0	1	0	0
Coffee time	2	0	97	1	0	0
Early morning	39	1	1	59	0	0
Clean up	1	1	0	0	96	2
Sandwich time	1	0	1	0	1	97

(u) KNN with primitive set 7

	Predicted Activities (%)					
True Activities	*Idle*	*Relaxing*	*Coffee time*	*Early morning*	*Clean up*	*Sandwich time*
Idle	93	3	1	1	1	1
Relaxing	51	48	0	1	0	0
Coffee time	2	0	97	1	0	0
Early morning	39	0	1	60	0	0
Clean up	1	0	0	0	97	2
Sandwich time	1	0	1	0	0	98

(v) BN with primitive set 8

	Predicted Activities (%)					
True Activities	*Idle*	*Relaxing*	*Coffee time*	*Early morning*	*Clean up*	*Sandwich time*
Idle	86	4	3	2	3	2
Relaxing	12	86	0	2	0	0
Coffee time	3	0	95	2	0	0
Early morning	22	1	0	77	0	0
Clean up	1	1	1	0	92	5
Sandwich time	1	0	2	0	1	96

(w) DT with primitive set 8

	Predicted Activities (%)					
True Activities	*Idle*	*Relaxing*	*Coffee time*	*Early morning*	*Clean up*	*Sandwich time*
Idle	91	2	2	2	2	1
Relaxing	8	89	0	3	0	0
Coffee time	2	0	97	1	0	0
Early morning	20	1	0	79	0	0
Clean up	1	1	0	0	95	3
Sandwich time	1	0	1	0	1	97

(x) KNN with primitive set 8

	Predicted Activities (%)					
True Activities	*Idle*	*Relaxing*	*Coffee time*	*Early morning*	*Clean up*	*Sandwich time*
Idle	93	2	1	2	1	1
Relaxing	10	88	0	2	0	0
Coffee time	2	0	97	1	0	0
Early morning	20	0	0	80	0	0
Clean up	1	0	1	0	96	2
Sandwich time	0	0	1	0	1	98

We use the data from the eight different action primitive sets described in Table 5. As evident in the bar charts in Figs 5–7 and confusion matrices in Table 6 all the three classifiers show almost the same performance to predict each activity with each of the action primitive sets. However, the choice of different action primitive sets have different affects on the value of true positive rate to predict an activity even with the same classifier. Here we discuss each activity and the effect of the different action primitive sets on its prediction with the classification algorithms.

Figures 5–7 show the value of true positive rate to predict the activity *Idle* with eight different action primitive sets and three different classifiers. While being *Idle*, the subject is not performing any activity.We observe in the case of all three classifiers that although action primitive set S2 shows best performance to predict this activity, other action primitive sets are also not far behind. This result is not surprising as when the subject is idle, he/she is neither interacting with any object nor making any movement. Consequently, most of the time action primitives have null value during this activity. So action primitives extracted from any particular type of sensing modalities are not particularly significant in the case of recognizing this activity. Action primitives extracted from any sensing modality can easily detect whether the subject is idle or not.

Figures 5–7 show the value of true positive rate to predict the activity *Relaxing* with the eight different action primitive sets and three different classifiers. This activity proved to be the most difficult one to recognize. During this activity, the subject went outside the studio apartment to have a short walk. Looking at the value of true positive rate to predict this activity by different classifiers, it can clearly be seen that the classifiers perform a lot better using action primitive sets

S1, S2, and S3 than using action primitive sets S4, S5, S6, S7, and S8. Primitive set S1 contains action primitives that have been extracted from all sensor modalities. Primitive set S2 contains action primitives that have been extracted from wearable sensors worn on the subject body and limbs. Primitive set S3 contains action primitives that are extracted from wearable sensors worn on the human body that present action primitives about subject body locomotion, such as walking, siting, lying. Human body locomotion action primitives in all the three sets, i.e., S1, S2, and S3, exhibit better performance to recognize the *Relaxing* activity.

The main reason is that during this activity the subject is not interacting with any object. So most of the time action primitives extracted from sensors embedded in objects have a null value. This is the reason why action primitive sets from those sensors are not very successful to recognize this activity. The classifiers get almost the same feature for this activity as for the *Idle* activity. Comparatively, the high number of *Idle* activity overwhelmed the classifiers decision and classifiers got confused making a distinction between *Idle* and *Relaxing*. The classifiers detected most of the *Relaxing* activities as the *Idle* activity. Primitive sets S1, S2, and S3 show better performance because those sets include action primitives, such as walking and standing, extracted from wearable sensors worn on subject body. But when action primitive sets exclude action primitives extracted from wearable sensors worn on the subject body recognizing the *Relaxing* activity becomes very difficult. Consequently, wearable sensors provides information about human locomotion action primitives proved vital for recognizing this activity.

Figures 5–7 show the value of true positive rate to predict the activity *Coffee Time* with eight different action primitive sets and three different classifiers. During *Coffee Time*, the subject prepared coffee with milk and sugar by using a machine, took sips of coffee, and also interacted with different objects in the environment. As evident from the activity description this activity is more distinctive on the basis of objects that are used as the human action primitives. Subsequently, action primitive sets that contain information about object usage and arm movements show better performance.

Classifiers show almost the same performance to predict *Coffee Time* with action primitive set S1, S2, and S4. S1 contains action primitives extracted from all sensors described in Table 2. S2 contains action primitives extracted from the sensors worn on body and arms, while S4 contains the action primitive sets extracted from sensor worn on arms only. As action primitive set S3 that contains primitives extracted from sensors worn on body only could not show comparable performance, we conclude that arm sensors are very good to predict such activities. Right arm motion action primitives show better performance than the subject body motion action primitives and left arm motion primitives as shown by the values of true positive rate for action primitive sets S2, S5, and S6.

Right arm motion action primitives also show better performance than the object usage action primitives, as shown by the value of true positive rate for action primitive sets S5 and S7. Better performance of right arm motion primitives is due to the fact that most of the time the subject is interacting with objects with his/her dominant limb. In our opinion, if we have rich wearable sensors that can provide human locomotion primitives like sip, wearable sensors can considerably increase the probability of correctly predicting this activity.

Figures 5–7 show the value of the true positive rate to predict the activity *Early Morning* with eight different action primitive sets and three different classifiers. During this activity, the subject moved in the room and randomly checked some objects in the drawers and on the shelf. Although primitive sets S4, S5, S6, and S7 showed better performance in recognizing this activity as compared to recognizing the *Relaxing* activity, action primitive set S2 that contains action primitives extracted from wearable sensors worn on the subject body and limbs exhibited the better performance in this case.

The reason for their better performance was that during this activity the subject spends a lot of time in physical activities. Again, in this case, he/she has not interacted with objects available in the environment for much time. Object sensors had been able to recognize this activity when the subject had not interacted with some of the objects. Action primitives extracted from wearable sensors worn on human limbs are also not very helpful in recognizing the *Early Morning* activity as these sensors also provide information about human interaction with objects available in the environment. Wearable sensor providing human locomotion primitives again proved vital in this case.

Figures 5–7 show the value of true positive rate to predict the activity *Clean UP* with the eight different action primitive sets and three different classifiers. During this activity, the subject puts all objects used in their original places or the dish washer and cleans up the table. As compared with other action primitives,

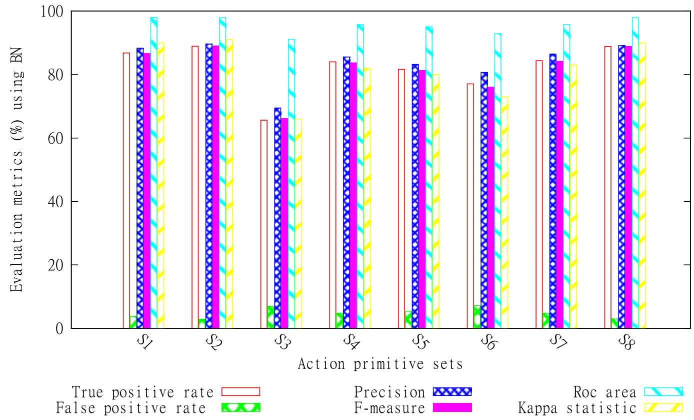


Fig. 8. Evaluation metrics showing the performance of Bayes Net to correctly predict ADL using different primitive sets.

body motion primitives, such as walk, sit, and stand, failed when those action primitives have been used alone to detect the *Clean Up* activity, as evident from the value of the true positive rate for S3. However, limbs locomotion primitives, such as reach, move, and release, showed comparatively better performance for the sensors used with right arm than the sensors used with left arm. If we compare the performance of all action primitive sets when these action primitive sets have been used alone, action primitives extracted from object usage sensors and action primitives extracted from sensors worn on right limbs give the best performance. Overall, the primitive set that used a combination of human locomotion primitives, limbs locomotion primitives, and object usage primitives showed the best performance in detecting the *Clean Up* activity.

Figures 5–7 show the value of true positive rate to predict activity *Sandwich Time* with eight different action primitive sets and three different classifiers. During this activity the subject interacted with different objects in the environment like bread, cheese, and salami, and had also used bread cutters, various kind of knives, and plates to prepare the sandwich. Later the subject ate that sandwich. Contrasting with the *Idle* activity when the subject was motionless most of the time and interacted with few objects, in this activity the subject has not only performed many low level physical activities, like cutting the bread, but has also interacted with various objects in the environment. As a result, all primitive sets performed well in the case of this activity compared with other activities. The action primitive set S4 that contains arm motion primitives shows better performance than action primitive set S7 that contains object usage action primitives. Action primitive set S4 also shows comparable performance with action primitive sets S1 and S2 that are the super set of S4. These results show that arm motions action primitives are most important type of action primitives to correctly predict this activity.

5.3. *Analyzing the impact of action primitives on the collective activity recognition*

This section presents the prediction performance of the classifiers for all the activities using each action primitive set. As we have seen in the previous section that all the classifiers show the same behavior with different action primitive sets, here in spite of reporting the identical results with different classifiers, we present our findings for a single classifier. We arbitrarily selected Bayes net classifier for that purpose. Figure 8 shows the weighted average of the true positive rate, false positive rate, precision, recall, f-measure, and kappa statistical measure of all activities classified using Bayes net classification algorithm. We added kappa statistical values to depict the measure of agreement (between the classifier using a particular action primitive set and the ground truth) normalized for chance agreement. Kappa statistic values are shown in Fig. 8 as the percentage measure. Kappa value of 100% indicate perfect agreement and Kappa value of 0% indicate chance agreement.

The action primitive set S1 includes action primitives extracted from all the sensing modalities available in the data set and the action primitive set S2 includes all the action primitive sets extracted from wearable sensors show the better overall performance. As S1 is the super set of all other sets, it is quite obvious that S1 shows better performance over other sets. Although S2 is a subset of S1 and contains action primitives extracted from wearable sensors only, classifiers are as successful in predicting activities using S2 as using S1. The reasons for the good performance of action primitive set S2 is the comprehensive nature of the action primitives extracted from wearable sensors. Wearable sensors not only provide information about the body motions like walk, sit, and stand but also indicate the usage of an object available in the environment. These primitives also proved very helpful in recognizing activities like *Idle*, when the subject is not performing any activity, *Early Morning*, when the subject is walking around and handling different objects, and *Relaxing* when the subject is sitting or lying. Kappa statistics values of 90% and 91% for the classifier using action primitive set S1 and S2 also depict the strong agreement between activities recognized by the classifier using these action primitive sets.

The action primitive set S3 consists of only body motion action primitives. Although this action primitive set is very good in predicting activities such as *Idle* and *Early Morning*, it was not helpful in predicting activities that involve a higher number of interactions with the environment and objects. Action primitive sets S5, S6, and S7 show the same overall performance. The classifiers show a slightly better performance in correctly predicting the activities by using S4 that contains arm motion primitives from both arms than S7 that contains object and environment usage action primitives. This means that arm motion action primitives can also be used in parallel with object and environment usage action primitives to predict human activities of daily living inside a home. Clearly primitive set S2 that used a combination of body motion and arm motion action primitives and S1 that is the super set of all action primitive sets show better performance. However, when we look at different action primitive sets individually the arms motion action primitive sets, especially the right arm action primitives, shows the best performance. The fact is also evident from the true positive rate of 81 and 77 percent and kappa statistical values of 80 and 74 percent for the Bayes net classifier to truly predict the activities using action primitive sets S5 and S6. Table 4 shows the detail of these action primitives and Table 5 shows the detail of sensing systems used to extract those action primitives.

6. Discussion and recommendations

In our experiments, we used the sliding window technique to extract feature histograms over the action primitives temporal data to use with classification algorithms to correctly predict human activities of daily living in a smart home. We used different jump sizes varying from one second to two minutes. We found that increasing the window size for the feature extraction, improves the performance of the classifiers to correctly predict human activity. However, increasing the window size beyond a certain limit (e.g. 60 seconds) does not necessarily reflect a better performance. So, we conclude that it is a very important step in designing a human activity recognition system to successfully decide the proper window size to extract feature histograms from temporal signal. Window size may also depend upon the requirements of the system that uses the prediction of human activity.

We used different classification algorithm to correctly predict human activities of daily living in a smart home. We used these classification algorithms with different set of action primitives containing different combination of object usage, arms motion, and body motion action primitives. The classifiers showed varying performance to predict human activities of daily living in smart home with different sets of action primitives. Spotting these action primitives needed different kind of sensing systems placed at different positions, e.g., spotting arms motion action primitives needed accelerometers worn on human arm. These facts show that instead of sticking with few data sets using specific kinds of sensing systems to predict human activities and working to improve the powers of classification algorithms to improve their performance, there is also a need to collect new data sets that use different sensing systems to collect data covering different sensing modalities.

We measured the performance of the classifiers to predict human activities in a smart home separately and collectively. We found that evaluation metrics showing the performance of classifiers to predict all activities together did not give enough information to analyze a human activity recognition system. Different classifiers and action primitive sets that may show better overall performance to correctly predict the human activities may have completely different performance when we look at their performance in detail. For example, as shown in Fig. 8 classifiers show better performance predicting ADL using action primitive set *S7* than using action primitive set *S3*. But if we look the classifier performance in predicting each activity individually we find that classifiers show a lot better performance to predict activities *Relaxing* and

Early Morning using action primitive set *S3* than using action primitive set *S7* as shown in Figs 5–7 and Table 6.

Considering the results of experiments we observe that the arms motion and the body motion action primitives that have been ignored in recognizing human activities in smart home environments play a significant role in improving the performance of human activity recognition systems. Even when the different types of action primitives have been used alone, the arm motion action primitives showed comparable performance to the object and environmental usage action primitives. The classifiers showed better performance using action primitive set *S2* that include all action primitives possible to extract from wearable sensors than *S7* that include all action primitives possible to extract from object and environmental sensors as shown in Figs 5–8. Body motion action primitives proved vital in recognizing human activities that do not involve much interaction with environment and objects, such as *Relaxing*, *Idle*, and *Early Morning*. Body motion and arm motion action primitives can also improve the performance of the classifiers to recognize other ADL of the same nature in a smart home, such as sleep, wake up, wash hands or watch television.

Object and environmental usage primitives completely failed in recognizing these activities. Limbs motion primitives, like reach, cut, and spread, also proved significant in recognizing those activities that include not only using but also performing actions on different objects. Examples of the actions performed on objects include cutting bread, applying bread spreads. Wearable sensors that can be used to extract action primitives like sip or bite are also important for correctly distinguishing activities like drinking coffee or eating a sandwich. Object or environmental usage sensors are very important to install in areas, such as kitchen, where human are expected to have greater interaction with those objects. Wearable sensors are significant in recognizing activities during which humans do not interact much with the environment, such as *Relaxing*. Sensors used with dominant limbs are more reliable in recognizing human activities than sensors used with other limbs. As arm and body motion action primitives in combination with object usage action primitives clearly outperformed only object usage action primitives in the recognition of all activities, it is indispensable to use wearable sensors to extract arms and body motion action primitives in smart environments to improve the performance of human activity recognition systems.

7. Conclusion

In this paper, we use object usage, body motion, and arms motion action primitives with different classification algorithms to recognize human activities of daily living in an ambient assisted living environment. The aim of this work is to analyze the impact of the aforementioned action primitives in the recognition of the human activities. During our experiments, we find that body and arm motion action primitives proved vital factor in predicting some of the activities, such as *Idle*, *Relaxing*, and *Early Morning* activities. While performing these activities human does not interact much with environment and objects. Consequently, the classifiers show considerably better performance in recognizing these activities using the combination of the body and arm motion action primitives than using the object and environment usage action primitives. The classifiers show better performance using right arm action primitives than using left arm action primitives and show almost the same performance while using action primitive sets containing only arm motion action primitives and object and environment usage action primitives. Consequently, we recommend to use wearable sensors that can detect human body and arms motion action primitives. We also recommend to use wearable sensors on the dominant limbs to improve the performance of the classifiers to predict ADL.

For future work, we plan to compare the effectiveness of different sensing systems that can be used to extract same action primitives in different home settings. We also plan to use the knowledge gained in this paper to design an autonomic human activity recognition system. This system will be able to dynamically select the most suitable classification techniques and opportunistically connect to the most appropriate set of deployed sensors to efficiently predict human activities in different conditions and home settings with optimal use of computing resources.

Acknowledgements

This work is partially supported by EU FP7 STREP Project SM4ALL (Smart hoMes for ALL) under Grant No. 224332, Science Foundation Ireland under the Principal Investigator research program 10/IN.1/I2980 Self-organizing Architectures for Autonomic Management of Smart Cities, and EU FP7 FET Project OPPORTUNITY under Grant No. 225938. We thank Claudia Villalonga for sharing her thoughts on an initial version of the paper.

References

[1] J.K. Aggarwal and M.S. Ryoo, Human activity analysis: A review, *ACM Comput. Surv.* **43**(3) (April 2011), 16:1–16:43.

[2] D.W. Aha, D. Kibler, and M.K. Albert. *Instance-Based Learning Algorithms*. Machine Learning, vol. 6, Kluwer Academic Publishers, 1991, pp. 37–66.

[3] A. Avci, S. Bosch, M. Marin-Perianu, R. Marin-Perianu, and P. Havinga, Activity recognition using inertial sensing for healthcare, wellbeing and sports applications: A survey, in: *Proc. of 23rd International Conference on Architecture of Computing Systems (ARCS)*, 2010, pp. 1–10.

[4] L. Bao and S.S. Intille, Activity recognition from user-annotated acceleration data, in: *Pervasive Computing*, Springer, Berlin Heidelberg, 2004, pp. 1–17.

[5] R.R. Bouckaert, Bayesian network classifiers in weka, Technical Report, Department of Computer Science, Waikato University, Hamilton, NZ 2005.

[6] T. Choudhury, G. Borriello, S. Consolvo, D. Haehnel, B. Harrison, B. Hemingway, J. Hightower, P. Klasnja, K. Koscher, A. LaMarca, J.A. Landay, L. LeGrand, J. Lester, A. Rahimi, A. Rea, and D. Wyatt, The mobile sensing platform: An embedded activity recognition system, *Pervasive Computing*, IEEE, **7**(2) (April 2008), 32–41.

[7] Y. Chu, Y.C. Song, R. Levinson, H. Kautz, Interactive activity recognition and prompting to assist people with cognitive disabilities, *Journal of Ambient Intelligence and Smart Environments*, IOS Press **4**(5) (2012), 443–459.

[8] D.J. Cook, Learning setting-generalized activity models for smart spaces, *Intelligent Systems*, IEEE **27**(1) (Jan.–Feb. 2012), 32–38.

[9] D.J. Cook, J.C. Augusto, and V.R. Jakkula, Ambient intelligence: Technologies, applications, and opportunities, *Pervasive and Mobile Computing*, Elsevier **5**(4) (2009), 277–298.

[10] D.J. Cook and M. Schmitter-Edgecombe, Assessing the quality of activities in a smart environment, *Methods of Information in Medicine* **48**(5) (2009), 480–485.

[11] J.A. À-Gracìa, P. Barsocchi, S. Chessa, and D. Salvi, Evaluation of localization and activity recognition systems for ambient assisted living: The experience of the 2012 EvAAL competition, *Journal of Ambient Intelligence and Smart Environments*, IOS Press **5**(1) (2013), 119–132.

[12] E. Gilman, O. Davidyuk, X. Su, and J. Riekki, Towards interactive smart spaces, *Journal of Ambient Intelligence and Smart Environments*, IOS Press **5**(1) (2013), 5–22.

[13] T. Gu, L. Wang, Z. Wu, X. Tao, and J. Lu, A pattern mining approach to sensor-based human activity recognition, *IEEE Transactions on Knowledge and Data Engineering* **23** (2011), 1359–1372.

[14] M.A. Hall, Correlation-based feature subset selection for machine learning, PhD thesis, University of Waikato, Hamilton, New Zealand, 1998.

[15] D.T.G. Huynh, Human activity recognition with wearable sensors, PhD thesis, TU Darmstadt, September 2008.

[16] S. Intille, K. Larson, E. Tapia, J. Beaudin, P. Kaushik, J. Nawyn, and R. Rockinson, Using a live-in laboratory for ubiquitous computing research, in: *Pervasive Computing*, Springer, Berlin Heidelberg, 2006, pp. 349–365.

[17] M.-W. Lee, A.M. Khan, and T.-S. Kim, A single tri-axial accelerometer-based real-time personal life log system capable of human activity recognition and exercise information generation, *Personal and Ubiquitous Computing* **15**(8) (2011), 887–898.

[18] B. Lepri, N. Mana, A. Cappelletti, F. Pianesi, and M. Zancanaro, What is happening now? Detection of activities of daily living from simple visual features, *Personal and Ubiquitous Computing* (2010), 1–18.

[19] B. Logan, J. Healey, M. Philipose, E.M. Tapia, and S. Intille, A long-term evaluation of sensing modalities for activity recognition, in: *UbiComp 2007: Ubiquitous Computing*, Springer, Berlin Heidelberg, 2007, pp. 483–500.

[20] T. Maekawa, Y. Yanagisawa, Y. Kishino, K. Ishiguro, K. Kamei, Y. Sakurai, and T. Okadome, Object-based activity recognition with heterogeneous sensors on wrist, in: *Pervasive Computing*, Springer, Berlin Heidelberg, 2010, pp. 246–264.

[21] A. Manzoor, H.-L. Truong, C. Dorn, and S. Dustdar, Service–centric inference and utilization of confidence on context, in: *Services Computing Conference (APSCC), 2010 IEEE Asia-Pacific*, IEEE, 2010, pp. 11–18.

[22] A. Manzoor, C. Villalonga, A. Calatroni, H.-L. Truong, D. Roggen, S. Dustdar, and G. Tröster, Identifying important action primitives for high level activity recognition, in: *Smart Sensing and Context*, P. Lukowicz, K. Kunze, and G. Kortuem, eds, Lecture Notes in Computer Science, Vol. 6446, Springer, Berlin/Heidelberg, 2010, pp. 149–162.

[23] S. McKeever, J. Ye, L. Coyle, and S. Dobson, Using Dempster-Shafer theory of evidence for situation inference, in: *Smart Sensing and Context*, Springer, Berlin Heidelberg, 2009, pp. 149–162.

[24] J.R. Quinlan, *C4.5: Programs for Machine Learning*, Morgan Kaufmann Publishers, San Mateo, CA, 1993.

[25] P. Rashidi, D.J. Cook, L.B. Holder, and M. Schmitter-Edgecombe, Discovering activities to recognize and track in a smart environment, *IEEE Transactions on Knowledge and Data Engineering* **23**(4) (April 2011), 527–539.

[26] N. Ravi, N. Dandekar, P. Mysore, and M.L. Littman, Activity recognition from accelerometer data, in: *Proc. of the National Conference on Artificial Intelligence*, Vol. 20, no. 3, Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999, 2005, p. 1541.

[27] D. Roggen, A. Calatroni, M. Rossi, T. Holleczek, K. Förster, G. Tröster, P. Lukowicz, D. Bannach, G. Pirkl, A. Ferscha, J. Doppler, C. Holzmann, M. Kurz, G. Holl, R. Chavarriaga, M. Creatura, and J. del R. Millán, Collecting complex activity data sets in highly rich networked sensor environments, in: *Networked Sensing Systems (INSS), 2010 Seventh International Conference on*, IEEE, 2010, pp. 233–240.

[28] E.M. Tapia, S.S. Intille, and K. Larson, *Activity Recognition in the Home Using Simple and Ubiquitous Sensors*. Springer, Berlin Heidelberg, 2004.

[29] P. Turaga, R. Chellapa, V.S. Subrahmanian, and O. Udrea. Machine recognition of human activities: A survey, *IEEE Transactions on Circuits and Systems for Video Technology* **18** (2008), 1473–1488.

[30] T. van Kasteren, G. Englebienne, and B. Kröuse, Transferring knowledge of activity recognition across sensor networks, in: *Pervasive Computing*, Springer, Berlin Heidelberg, 2010, pp. 283–300.

[31] T. van Kasteren, A. Noulas, G. Englebienne, and B. Kröse, Accurate activity recognition in a home setting, in: *Proc. of the 10th International Conference on Ubiquitous Computing*, ACM, New York, NY, USA, 2008, pp. 1–9.

[32] A.J. Viera and J.M. Garrett, Understanding interobserver agreement: The kappa statistic, *Fam. Med.* **37**(5) (2005), 360–363.

[33] L. Wang, T. Gu, X. Tao, H. Chen, and J. Lu, Recognizing multi-user activities using wearable sensors in a smart home, *Pervasive Mob. Comput.* **7**(3) (June 2011), 287–298.

[34] J. Wu, A. Osuntogun, T. Choudhury, M. Philipose, and J.M. Rehg, A scalable approach to activity recognition based on object use, in: *Proc. of the 11th International Conference on Computer Vision (ICCV)*, IEEE, Rio de, 2007.

[35] J. Ye, L. Coyle, S. Dobson, and P. Nixon, Using situation lattices in sensor analysis, in: *IEEE International Conference on Pervasive Computing and Communication*, IEEE, 2009, pp. 1–11.

[36] J. Ye and S. Dobson, Exploring semantics in activity recognition using context lattices, *Journal of Ambient Intelligence and Smart Environments*, IOS Press **2**(4) (2010), 389–407.

[37] P. Zappi, C. Lombriser, T. Stiefmeier, E. Farella, D. Roggen, L. Benini, and G. Tröster, Activity recognition from on-body sensors: Accuracy power trade-off by dynamic sensor selection, in: *Proc. of the 5th European Conference on Wireless Sensor Networks*, IEEE, New York, 2008, pp. 17–33.