Expert Systems with Applications 40 (2013) 1662–1674

Contents lists available at SciVerse ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Elderly activities recognition and classification for applications in assisted living

Saisakul Chernbumroong [a,b], Shuang Cang [c], Anthony Atkins [a], Hongnian Yu [d,*]

[a] Faculty of Computing, Engineering and Technology, Staffordshire University, Stafford ST18 0AD, UK
[b] Department of Software Engineering, College of Arts, Media and Technology, Chiang Mai University, Chiang Mai 50000, Thailand
[c] School of Tourism, Bournemouth University, Poole, Dorset BH12 5BB, UK
[d] School of Design, Engineering & Computing, Bournemouth University, Poole, Dorset BH12 5BB, UK

ARTICLE INFO

Keywords:
Assisted living systems
Activities of Daily Livings (ADLs)
Wrist-worn multi-sensors
Elderly care
Feature selection and classification

ABSTRACT

Assisted living systems can help support elderly persons with their daily activities in order to help them maintain healthy and safety while living independently. However, most current systems are ineffective in actual situation, difficult to use and have a low acceptance rate. There is a need for an assisted living solution to become intelligent and also practical issues such as user acceptance and usability need to be resolved in order to truly assist elderly people. Small, inexpensive and low-powered consumption sensors are now available which can be used in assisted living applications to provide sensitive and responsive services based on users current environments and situations. This paper aims to address the issue of how to develop an activity recognition method for a practical assisted living system in term of user acceptance, privacy (non-visual) and cost. The paper proposes an activity recognition and classification method for detection of Activities of Daily Livings (ADLs) of an elderly person using small, low-cost, non-intrusive non-stigmatize wrist worn sensors. Experimental results demonstrate that the proposed method can achieve a high classification rate (>90%). Statistical tests are employed to support this high classification rate of the proposed method. Also, we prove that by combining data from temperature sensor and/or altimeter with accelerometer, classification accuracy can be improved.

1. Introduction

In 2011 there were approximately 647 million people aged 65 years and over in the world and is expected to accelerate to 1.91 billion people in 2050 (United Nations & Social Affairs, 2011). Population ageing phenomenon has affected every human in many ways, especially in healthcare as the health of elderly persons deteriorate with increasing ages. Problems such as a rise in expenditure of healthcare, quality of care e.g. increased burden on caregivers, insufficient and inefficient care are more likely to happen. Majority of elderly people prefer to remain in their own homes for as long as possible (Bayer, Harper, & Greenwald, 2000). Also, the cost of care home can be expensive comparing to assisted living facilities. Currently, there are systems which facilitate self-care and extend the independence of ageing population. Such systems are often known as assisted living systems. The main aims of the systems are to enable elderly people to independently live longer in their own homes, to enhance living qualities and to reduce costs for society and public health systems (Kleinberger, Becker, Ras, Holzinger, & Muller, 2007). Assisted living systems can help support elderly persons with their daily activities in order to help them maintain healthy and safety while living independently. At the present, there are a number of off-the-shelve products available in the market e.g. fall monitoring system on mobile phone, emergency alarm, etc. Usually they are closed, stand-alone systems with limited ability to describe actual situations, often too difficult for elderly people to use and ineffective in emergency situations (Kleinberger et al., 2007). Clearly, there is a need for the assisted living solution to become intelligent and also practical issues such as user acceptance and usability need to be resolved in order to truly assist elderly people.

Small, inexpensive and low-powered consumption sensors are now available which can be applied in healthcare in order to promote and compromise opportunities and challenges due to population ageing. In assisted living applications, such sensors can also be used to provide sensitive and responsive services based on users current environments and situations. This study aims to develop an activity detection method which could be used to provide adaptive services for users. Recognition of ADLs which are activities for taking care of oneself in daily life is of interest as it would allow intelligent and adaptive assistive services for elderly people. The study investigates the research question of how to develop an activity recognition method for assisted living systems that can achieve good accuracy while taken practicality issues including user acceptance, privacy (non-visual) and cost into account. It is believed that these are vital keys toward the intelligent

* Corresponding author.
E-mail address: yuh@bournemouth.ac.uk (H. Yu).

0957-4174/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.eswa.2012.09.004

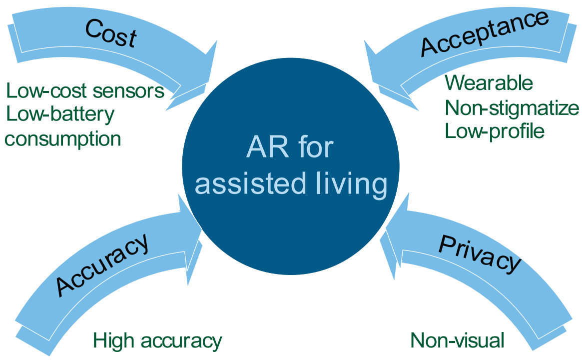

Fig. 1. Activity recognition for a practical assisted living system.

assisted living solution especially for elderly people who are the most important stakeholders. Fig. 1 illustrates the factors required for activity recognition in a practical assisted living system.

The main contributions of the paper are:

- Propose an activity recognition method for detection of ADLs of an elderly person using small, low-cost, non-intrusive non-stigmatize wrist-worn sensors. This method is based on data collected from a group of elderly people in a real living-home setting.
- Develop a feature selection method named feature combination which emphasizes on the performances of a combination of features rather than single feature. It uses forward selection to find the best combination of feature for each data set and monitor the network accuracy along in order that overlapped features are not selected.
- Prove two hypothesizes: H1: the proposed method can achieve high classification rate (>90% accuracy) which is higher than previous similar works (Fleury, Vacher, & Noury, 2010; Maurer, Rowe, Smailagic, & Siewiorek, 2006). H2: activity classifications accuracy can be improved by combining data from temperature sensor and/or altimeter with accelerometer.

The rest of this paper is organized as follows: Section 2 surveys the related works and addresses the problems in current systems. Section 3 describes the methodology used in the development of the activity recognition method including the choice of sensors and activities, feature selection and classifier construction. Section 4 presents experimentation and results of the proposed activity recognition method and hypotheses testing. Section 5 discusses the experimental results and explores how the proposed work may be used in assisted living systems to assist elderly people.

2. Assisted living systems

2.1. System requirement

Kleinberger et al. (2007) stated the following three major requirements for an assisted living system which need to be met in order to fulfill its purpose and potential to assist vulnerable people:

(1) *High acceptance*: System needs to be ambient and unobtrusive.
(2) *Adaptation*: Able to adapt to changing situations or abilities of the individual and environment to satisfy individual needs.
(3) *High usability*: Services must be provided in an accessible way.

These three characteristics can be viewed from two perspectives namely practical and technical aspects. The aim of the practical aspect is to satisfy users needs in term of practicality i.e. high usability and acceptance. Several studies on perceptions toward assisted living technology (Demiris et al., 2004, Demiris, Hensel, Skubic, & Rantz, 2008; Chernbumroong, Atkins, & Yu, 2010). have identified concerns arisen from the use of technology for elderly care such as user acceptance, system cost, usability and privacy issues. For example, a system which requires users to wear special equipment may be perceived as stigmatization or too complicated to use resulting in low acceptance. In a mobility aided system (Yu, Spenko, & Dubowsky, 2003) for example, the user interface is critical requirement as it has the direct physical interaction with the users. An interview-based investigation by Demiris et al. (2004, 2008) also showed that elderly people were concerned about privacy violation, visibility and accuracy of the assisted living systems. Even if systems could deliver the best services for assisting people unless they are easily accessible and usable and address the real needs and concerns of the users, they will not be accepted. The cost of assisted living systems is another important issue. For a practical solution in assisted living, the systems need to be cost-effective to make it affordable for general population especially elderly people who may be on benefits or pensions.

In the technical aspect, the aim is to enable assisted living systems to become intelligent in order to adaptively assist elderly people in changing and dynamic environment. Systems with such aims are also referred to as ambient intelligent systems or smart homes (Aarts, 2004). Ambient intelligence technology used in assisted living solutions can provide some hands-on and supports based on current user status to maintain elderly people independences, reducing the healthcare cost while increasing quality of life (Kleinberger et al., 2007). To offer intelligent services, the key is to understand users environments such as surrounding temperature and users activities etc. in order to provide adaptive assistance to users. In this study, we are interested in understanding the users current activities which is often referred to as activity recognition.

2.2. Activity recognition for an assisted living system

Many systems that are able to detect users activities have been proposed and developed for effective assisting services (Fleury et al., 2010; Hong, Kim, Ahn, & Kim, 2010, 2009; Maurer et al., 2006; Martínez-Pérez, González-Fraga, Cuevas-Tello, & Rodrıguez, 2012; Najafi et al., 2003; Zhu & Sheng, 2011). For example, Hong et al. (2010) developed a system that detects user activities which can be used to monitor energy expenditure and calorie intake. Hong et al. (2009) proposed a framework based on information management and evidence fusion to infer ADLs using anonymous binary sensors. Fleury et al. (2010) monitored surrounding environments and user activities in order to ensure the elderly living safely and independently in their own homes. These systems can detect a range of activities e.g. walking, feeding, using telephone, etc. of a user. In assisted living applications, detected activities can be divided into two categories namely ambulatory activities and ADLs. Ambulatory activities are activities that related to walking including static postures (e.g. standing, sitting), transition activities (e.g. sit-to-stand, stand-to-sit), and dynamic activities (e.g. walking). Examples of activity recognition systems which recognize ambulatory activities were reported in the literatures such as Najafi et al. (2003), Zhu and Sheng (2011) and Fuentes, Gonzalez-Abril, Angulo, and Ortega (2012). ADLs, on the other hand, cover a broader range of activities often found in daily living. Majority of activity recognition for assisted living systems were developed to detect ADLs such as feeding, watching TV, using telephone, cooking, etc. (Fleury et al., 2010; Hong et al., 2010, 2009; Martínez-Pérez et al., 2012).

Different types of sensors have been used in activity recognition systems for detecting elderly people ADLs such as accelerometer (Fleury et al., 2010; Hong et al., 2010, 2009; Maurer et al., 2006; Martınez-Pérez et al., 2012; Najafi et al., 2003; Zhu & Sheng, 2011; Banos, Damas, Pomares, Prieto, & Rojas, 2012; Fuentes et al., 2012), gyroscope (Najafi et al., 2003), Radio Frequency Identification (RFID) (Hong et al., 2010; Martínez-Pérez et al., 2012; Huang, Lee, Kuo, & Lee, 2010), binary sensor (Fleury et al., 2010; Hong et al., 2009), camera (Costa, Castillo, Novais, Fernndez-Caballero, & Simoes, 2012; Olivieri, Conde, & Sobrino, 2012; Martínez-Pérez et al., 2012; Zhu & Sheng, 2011), microphone (Fleury et al., 2010; Maurer et al., 2006), activity sensor (Kwon, Shim, & Lim, 2012; Botia, Villa, & Palma, 2012), etc. Although, prior research in activity recognition are often done via visual information (see a recent review in Chaaraoui, Climent-Prez, & Flrez-Revuelta (2012)), this may not be suitable for elderly care applications due to privacy violation. Previous studies (Demiris et al., 2004, 2008; Chernbumroong et al., 2010) shown that greater concerns were raised in the use of video sensor in assistive technology. Also, some of these sensors may be presented in special equipment e.g. RFID glove (Hong et al., 2010), PDA with RFID reader (Huang et al., 2010) which may be not easily accepted by the elderly and/or could incur high cost. The distribution of these sensors also varies in different systems. Sensor location can be divided into three main categories namely on-body (Fuentes et al., 2012; Maurer et al., 2006; Najafi et al., 2003), environment (Hong et al., 2009; Kwon et al., 2012; Botia et al., 2012) and combination of both (Fleury et al., 2010; Hong et al., 2010; Martínez-Pérez et al., 2012; Zhu & Sheng, 2011; Huang et al., 2010). The location of the sensor is linked with the acceptance level of an assisted living system. Certain sensors location or multiple sensor locations may prevent elderly people to perform activities normally or may cause discomfort. The numbers of sensors used in an activity recognition system also vary greatly. A system with the distributed sensors in environment normally requires a larger number of sensors. For example, studies such as Huang et al. (2010) and Hong et al. (2010) deployed RFID tags on numerous objects in homes. Similarly, Hong et al. (2009) required home objects e.g. cup, fridge, tea, etc. to be equipped with contact switch sensors. However, this approach may be time consuming and not feasible to set up a system.

A variety of activity classification methods have been proposed for detecting activities of elderly people. In term of features used for classification, the majority of the systems used heuristic features (Hong et al., 2010; Najafi et al., 2003; Zhu & Sheng, 2011). Usually features are chosen from literatures or directly from raw sensor data. Features selected from this method may not yield the optimum features or may consist of a large feature set which indicates the redundant features. Feature selection, on the other hand, assesses each feature in order that only important features are selected. In this work, a feature selection method is proposed which extends feature ranking i.e. Clamping (Wang, Jones, & Partridge, 2000) and subset selection i.e. forward selection techniques. The clamping technique uses feed-forward network and feature ranking is based on the effect of the absent of that feature in the network. The advantage of this technique is its simplicity allowing irrelevant and relevant features to be identified. This technique, however, only considers the performance of a single feature, also when use with forward selection, selected features may be highly overlapped. To overcome this limitation, we propose another technique which uses forward selection to find the best combination of features for a data set and monitor the network accuracy to avoid the selecting overlapped features.

In term of classifier development, different techniques have been used such as machine learning e.g. Support Vector Machine (SVM) (Fleury et al., 2010; Fuentes et al., 2012), Decision Tree (Chernbumroong, Atkins, & Yu, 2011; Maurer et al., 2006), Multi Perceptron (MLP) Neural Network (Györbıró, Fábián, & Hományi, 2009; Parkka et al., 2006), etc., hierarchy-based (Najafi et al., 2003), combination of techniques (Zhu & Sheng, 2011; Hong et al., 2010) and other technique e.g. rule-based (Martínez-Pérez et al., 2012; Botia et al., 2012), etc. Experimental studies were conducted on three different classification algorithms namely MLP, Radial Basis Function (RBF) and SVM. Also, it is noted that our study uses the data set collected from a group of elderly people. Many studies on activity recognition for assisted living for elderly people were conducted using data from young adults (Fleury et al., 2010; Zhu & Sheng, 2011) or from simulation which may fail to represent true activity characteristics of elderly people.

The accuracy of activity recognition for assisted living systems vary depending on the number of activities, type of activities, number of sensors used, sensor location, etc. Activity recognition models which detect ambulatory activities often achieve high accuracy over 88% (Najafi et al., 2003; Zhu & Sheng, 2011) while models that detect ADLs often achieve slightly lower accuracy. It should be noted that it is difficult to make a fair comparison between systems due to aforementioned variations. However, based on inputs and outputs of the models, four similar works are used for comparison. Firstly, Fleury et al. (2010) used multiple sensors inputs (on-body and environment) and SVM based classification to detect seven ADLs: sleeping, preparing and having meal, dressing/undressing, hygiene activities (wash hand and teeth), bowel movement, communication. Secondly, Maurer et al. (2006) used multi-sensor wrist-worn equipment to detect five ADLs including walking, walking upstairs, walking downstairs, sitting and running. Thirdly, Huang et al. (2010) proposed an activity recognition approach using pattern mining and matching techniques. In their system, a user carries a PDA with a RFID reader which can detect the touched objects in order to infer activities. They can detect four activities including brushing teeth, cooking, eating, and washing dish. Finally, Hong et al. (2010) used three triaxial accelerometers worn on thigh, waist and wrist and RFID to detect 5 body states and 16 hand activities.

3. Methodologies

3.1. Choice of activities of daily livings

The study aims to recognize ADLs that are activities necessary in taking care of oneself and commonly occur in real daily living. Recognized ADLs can be used for evaluating elderly independence. There are two main types of ADLs, namely basic ADLs (BADLs) and instrumental ADLs (IADLs). The BADLs are activities necessary for self-care while IADLs are not. Examples of BADLs are feeding, bathing, dressing, grooming, etc. IADLs are such as using telephone, house work and doing laundry, etc.

In this research, both BADLs and IADLs are of interest as they cover the majority of activities occurred in independent living situation. For the BADLs, five activities from the Barthel Index (Mahoney & Barthel, 1965) are selected namely feeding, grooming (brushing teeth), dressing, mobility (walking) and stairs. It is noted that activities that are not selected are due to the difficulty in data collection such as toilet use, bathing, etc. In addition, sleeping/lying down activity is also chosen as it is a common activity in everyday life. For IADLs, housework activities i.e. washing dishes, ironing and sweeping and leisure activities i.e. watching TV are selected. In total, 11 activities of BADLs and IADLs are of interest in this research. The activities and their descriptions are listed in Table 1.

3.2. Choice of sensors and sensor setup

A diversity of sensors has been embedded into wrist-worn equipment. This includes regular sensors such as accelerometers,

Table 1
Activities of Daily Livings studied in this study.

Type	Activities	Activity and independence description
BADLs	Feeding	Feeds self without assistance (using spoon and fork)
	Brushing teeth	Brushes self teeth without assistance, including the use of toothpaste
	Dressing	Gets clothes and dresses without any assistance except for tying shoes
	Walking	Walks from one place to another without assistance
	Walking upstairs	Walks up the stairs without assistance
	Walking downstairs	Walks down the stairs without assistance
	Sleeping/lie down	Sleeps or lies down on a bed
IADLs	Washing dishes	Washes dishes, glasses
	Ironing	Irons shirt, trousers, pillow case, etc
	Sweeping	Sweeps floor using broom
	Watching TV	Sits and watches television

gyroscope and magnetometer, etc. as well as bio-sensors such as electrode for measuring changes in skin conductance (Poh, 2011). The selection of wrist-worn sensors usually depends on the target application as different sensors provide different data which may be more useful in certain applications than others. Since the purpose of this study is to detect activities of daily living of an elderly person, it is sensible to use sensors which are sensitive to changes in activities and can reflect different types of activities well. Also, sensors should be small, low-cost, non-intrusive, non-stigmatize and low-powered to be suitable for continuous usage in elderly care applications. A study by Chernbumroong et al. (2011) has demonstrated that an accelerometer is capable of detecting human activities, thus it is selected. In this study, temperature sensor and altimeter are also selected as it is hypothesized that the combination of these three sensors can help achieve higher activity classification performance.

This work investigates three types of sensors, namely accelerometer, temperature sensor and altimeter. These sensors are integrated on a normal sport watch as shown in Fig. 2. The eZ430-Chronos watch is based on the CC430F6137 Microcontroller with the MSP430 CPU from Texas Instruments (2010). The watch has an integrated 868 MHz wireless transceiver which allows communication with the PC through a USB RF access point wirelessly. The watch also contains 8 KB of flash memory available for data logging. The on-board accelerometer can measure acceleration in 3 dimension at a range of up to $\pm 2G$ ($G = 9.81$ m/s^2) and a sampling rate of 33 Hz. The accelerometer actual sampling rate is 100 Hz, however, to reduce the energy consumption, the watch only transmits the third data set. The sample rate of temperature sensor and altimeter is 1 Hz.

3.3. The proposed activity recognition method

The proposed method for detection of ADLs of elderly using wrist-worn multi-sensors is depicted in Fig. 3 which can be divided into two main stages: offline and online stages. The offline stage shown in dashed box involves the processes required for activity recognition model development. The result from the offline stage is an activity recognition model which is used later in online or real-time activity recognition. The process of online activity recognition of an elderly person is shown in a solid box.

3.3.1. Multi-sensors inputs

Firstly, the model receives three sensor inputs i.e. acceleration, temperature and altitude which are wirelessly transmitted from the watch worn on the dominant wrist of the elderly person. The

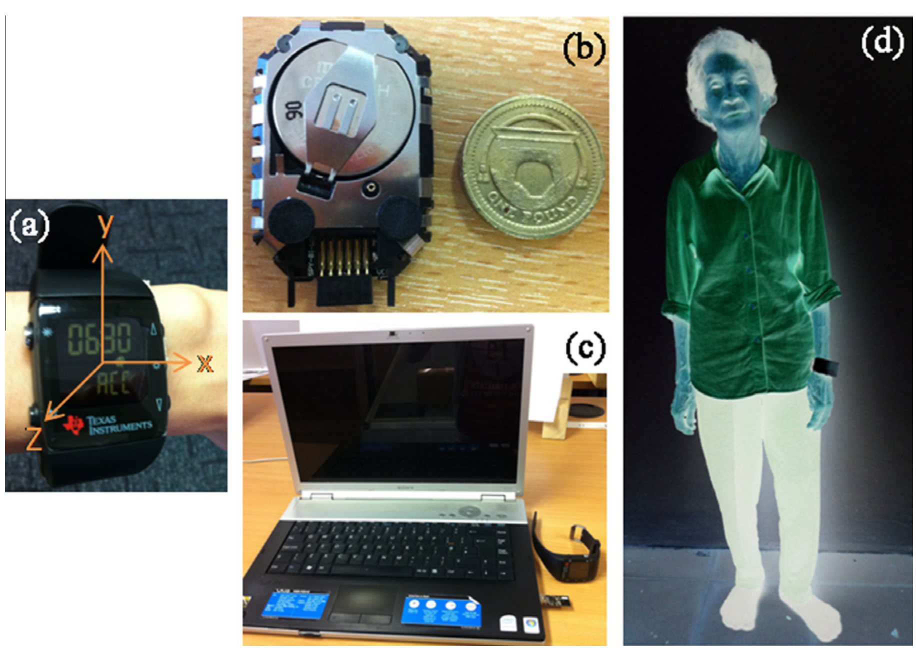

Fig. 2. Experiment equipments and setup. (a) Acceleration axis. (b) A watch module comparing to a pound coin. (c) Equipment used for experiments. (d) Sensor location on a participant.

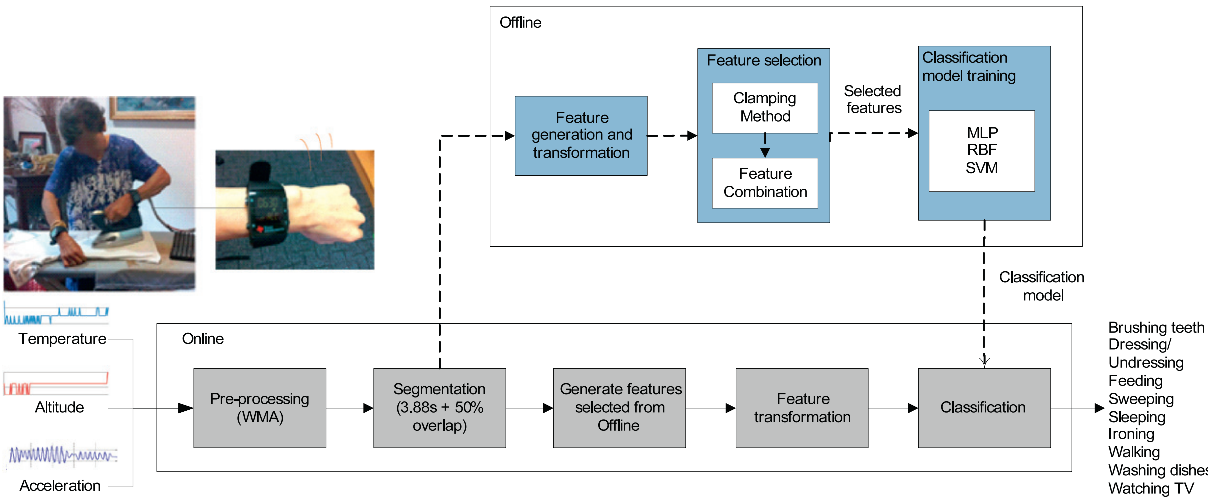

Fig. 3. The proposed activity recognition method to detect elderly ADLs using wrist-worn multi-sensors.

acceleration data *Acc* is consisted of acceleration from three axes (Fig. 2). The data received at any time t is:

$$data_t = \{Acc^{Ub}, Temp^{Ub}, Alt^{Ub}\}$$
$$Acc^{Ub} = \{Acc_x, Acc_y, Acc_z\}$$

where Ub denotes the sensor sample rate. For example in our application, Ub of the accelerometer is 33 Hz.

3.3.2. Pre-processing and segmentation

To smooth the graphs and remove the outlier, the acceleration data is pre-processed using the WMA technique. The WMA assigns different weights on data at different points, specifically higher weights are given to more recent data. For any set of n accelerations, acceleration at t is calculated as:

$$Acc_t = \sum_{i=1}^{n} w_i Acc_{t-i+1}$$
$$\sum_{i=1}^{n} w_i = 1$$

In order to prepare the input from the sensor data, the Sliding-window technique is used. This technique is commonly used for separating time series data into the input vector without losing information. An experiment on the different window length was carried out where it was decided to use window of 3.88 s (128 time frames). The sensor data from accelerometer, temperature sensor and altimeter were divided into 128-window length with 50% overlapping. For a window size l, the patterns are:

$$Patterns = \begin{bmatrix} 1 & 2 & \cdots & l \\ \frac{l}{2}+1 & \frac{l}{2}+2 & \cdots & \frac{l}{2}+l \\ \vdots & \vdots & \ddots & \vdots \\ \frac{i-1}{2}l+1 & \frac{i-1}{2}l+2 & \cdots & \frac{i-1}{2}l+l \end{bmatrix}$$

where:

$$0 \leqslant i \leqslant \frac{2 \times All\ data}{l} - 1$$

3.3.3. Feature generation and transformation

Experimentation demonstrated that 13 different features from both time and frequency domains were useful in human activity classification (Chernbumroong et al., 2011). In this work these features are also computed from the collected sensor data. However, the features are not only calculated from the acceleration magnitude $\sqrt{x^2+y^2+z^2}$, but also from raw accelerations of X, Y, Z axis, temperature and altitude as well. In addition, other features including maximum, root-mean-square (RMS), and main axis are also calculated. A list of features is displayed in Table 2 containing 12 features from X-axis acceleration, 12 features from Y-axis acceleration, 12 features from Z-axis acceleration, 10 features from acceleration magnitude, 8 features from temperature, 8 features from altitude and 1 feature from acceleration. The calculated features are then transformed into range. Scaling helps avoiding features with larger numeric ranges dominating features with smaller numeric ranges. It also reduces numerical difficulties during calculation (Hsu & Lin, 2002). In the MLP neural network which uses the gradient descent method i.e. backpropagation, scaling can help in faster convergence.

3.3.4. Feature selection

In order to understand the importance of each feature and eliminate the least promising one, it is essential to apply an effective feature selection technique. Also, a classifier with minimum inputs and optimum performance can be produced. The Clamping technique (Wang et al., 2000) is used to obtain the feature ranking where each feature is clamped to a fixed value (mean of each feature $\bar{x}$ was used) and the impact of the clamped network generalization performance, $g(X|x_i = \bar{x})$ to the network generalization performance, $g(X)$ is calculated using:

Table 2
Number of features calculated from each sensor data.

Sensor data	Time-domain features	Frequency-domain features
Acceleration X-axis, Acceleration Y-axis, Acceleration Z-axis, Acceleration magnitude, Temperature, Altitude	Mean, minimum, maximum, standard deviation, variance, range, root-mean-square, correlation, difference, main axis	Spectral energy, spectral entropy, key coefficient
Total number of features	45	18

$$Im_i = 1 - \frac{g(X|x_i = \bar{x})}{g(X)}$$

Clamping the most important feature will highly affect generalization performance while the redundant features will show no adverse effect. For an N data set, the rankings can be combined using the Borda Count technique which is a kind of plurality voting where each vote is associated with a specified ranking point. A given feature was associated with a point that is related to its importance. The most important feature is associated with the highest possible ranking value e.g. the total number of features. For example, in a feature space size n, feature X is the most important, hence it is associated with n points. Feature Y is the second most important thus, $n - 1$ points is associated, and so on. Finally, the final feature ranks can be obtained by sorting the summation of the points of each feature.

However, the feature ranking using the Clamping technique can only consider the performance of a single feature. Feature selection based on this ranking may discard the features which are useless in itself but help improve classification performance when combined with other features. Also, the features with high ranking may be overlapped with other high ranking features. To overcome this, a Feature Combination technique is proposed which emphasizes on the performances of a combination of features rather than single feature. The idea is to use forward selection to find the best combination of features for a data set. A feature is added to the lists by its importance and difference in accuracy is calculated along each addition. By monitoring the accuracy difference, the feature which is highly overlapped with already added features will not be included into the list. This technique also allows the weaker feature which is not overlapped with existed features to be selected.

Starting from an empty list, a feature is added according to its ranking. For any current feature list using p features, mean of accuracy $(\overline{M\_Acc_p})$ of validation set is calculated and compared with mean of accuracy $(\overline{M\_Acc_{p-1}})$ of the previous feature list i.e. using $p - 1$ features. If $(\overline{M\_Acc_p})$ is less than or equal to $(\overline{M\_Acc_{p-1}})$, then the recently added feature is removed from the list. This process is carried out until all features have been tested. For an N data set, results are combined using majority voting resulted in a new feature ranks. Fig. 4 describes the pseudo code of the feature selection.

```
PROCEDURE Feature Combination

TRAIN Network using All feature
COMPUTE Network generalized performance
FOR each Feature in All feature
    SET Input to All features
    SET Feature in Input to mean Feature
    TEST Network using Input
    COMPUTE Clamped network generalized performance
    COMPUTE Impact of clamped network generalized performance   (Im)
END LOOP
SORT Im By DESC.
SET Rank to Im

INITIALIZE List
FOR each Feature in Rank
    ADD Feature to List
    TRAIN Network using List
    TEST Network using List
    COMPUTE Accuracy_current
    IF Accuracy_current<=Accuracy_previous
        REMOVE Feature from List
    ELSE
        SET Accuracy_previous to Accuracy _current
    END
END LOOP
```

Fig. 4. Pseudo code of feature combination.

3.3.5. Neural networks and Support Vector Machine for classification

MLP exploits the idea of the nervous system in which numerous inputs are connected to numerous outputs. These connections are associated with weights and the outputs are usually calculated from activation functions, such as sigmoid functions, of summation of weighted inputs. MLP is capable of learning any nonlinear functions by adjusting the connection weights to minimize the error of the output. Several works on sensor-based activity classification have been conducted using MLP (Chernbumroong et al., 2011; Györbı́ró et al., 2009; Parkka et al., 2006; Zhu & Sheng, 2011). The RBF network is another type of neural network which uses a radial basis function as an activation function. Also, instead of using summation of weighted inputs for an activation function input, the distant between its weights and input is used.

A limited number of activity recognition studies were carried out using SVM (Fleury et al., 2010; Fuentes et al., 2012). The main concept of this classifier is to find decision boundaries which separate the data with the largest margin as possible. SVM maps an input into a new higher dimensional space by some kernel functions such as linear kernel, Gaussian kernel and polynomial kernel, etc. It then finds a hyperplane with maximal margin to separate the data. The advantages of SVM are that it can produce a global optimal solution and work well on small data set (Hsu & Lin, 2002; Hsu, Chang, & Lin, 2003).

For K class classification, let multi-sensors $X = \{x_1, x_2, \ldots, x_n\}$ be the input where N is a sample size. Given a training example (X_i, y_i) where y_i is the class variable which has K values. SVM projects the training example into a new higher dimensional space using the radial Gaussian kernel function below:

$$exp(-\gamma \|x_i - x_j\|^2), \quad \gamma > 0$$

where $\gamma = \frac{1}{(2\sigma^2)}$ and σ represents the width of the Gaussian kernel. For m training examples, SVM finds a solution which satisfies following optimization objective (Hsu et al., 2003):

$$\min_{w,b,\xi} \frac{1}{2} w^T w + C \sum_{i=1}^{m} \xi_i$$

subject to:

$$y_i(w^T f(x_i) + b) \geqslant 1 - \xi_i$$
$$\xi_i \geqslant 0$$

The slack term ξ_i is used to relax the constraints allowing misclassified examples. The associated cost parameter C is used for penalizing ξ_i. Generally, SVM is designed for classification of binary problems. For our K-class problem, one-VS-one approach was applied. This method achieves a good classification result and uses less computing time which is suitable for practical applications (Hsu & Lin, 2002). In one-VS-one approach, $\frac{(K \times (K-1))}{2}$ binary classifiers used for classifying class C_i and C_j are built. For all training examples from class i and j, SVM solves the following optimization problem (Hsu & Lin, 2002):

$$\min_{w^{ij}, b^{ij}, \xi^{ij}} \frac{1}{2} (w^{ij})^T w^{ij} + C \sum_t \xi_t^{ij}$$

Subject to:

$$(w^{ij})^T f(x_i) + b^{ij} \geqslant 1 - \xi_t^{ij} \quad \text{if } y_t = i$$
$$(w^{ij})^T f(x_i) + b^{ij} \leqslant -1 + \xi_t^{ij} \quad \text{if } y_t = i$$
$$\xi_i \geqslant 0$$

3.4. Performance measurement

In this study, four measurements are used to measure the method performance. First, accuracy is the ratio that the model correctly

Table 3
Confusion matrix.

		Predict	
		+	–
Actual	+	True positive	False negative
	–	False positive	True negative

classifies the input. It is commonly used measurement for classification and can be calculated using the equation $accuracy = \frac{TruePositive+TrueNegative}{AllInstances}$. Next are precision and recall which focus more on the number of false positives and false negatives. Precision is a measurement indicating the fraction that the model correctly predicts true class while recall indicates the fraction that the model detects true class. Precision and recall can be obtained using equations $precision = \frac{TruePositive}{TruePositive+FalsePositive}$ and $recall = \frac{TruePositive}{TruePositive+FalsePositive}$, respectively. Finally, F-score is an extended measurement of accuracy. While accuracy ignores the false positive results which means that it cannot differentiate if the classifier is being discriminative to a particular class, F-score does not. F-score equally combines precision and recall. F-score can be computed using equation $F-score = \frac{2\ recall\ precision}{recall+precision}$. The relationship between true positive, true negative, false positive and false negative instances are displayed in Table 3.

3.5. Hypothesis testing

3.5.1. H1: The proposed method can achieve more than 90% accuracy

Firstly, the accuracies of the proposed activity recognition method are tested for normality using the Shapiro–Wilk test. One-tailed one-Sample T-test statistics at 5% significance level is used to test the hypothesis that the mean accuracy of the proposed method is higher than 90% (H_0:$\mu \leqslant 90\%$, H_a:$\mu > 90\%$).

3.5.2. H2: Classification accuracy can be improved by combining data from temperature sensor and/or altimeter with accelerometer

To test the hypothesis that using the combination of accelerometer, temperature sensor and altimeter would improve the activity classification performance, four different configurations are used: (A) features from only accelerometer, (B) accelerometer with temperature sensor, (C) accelerometer with altimeter and (D) combination of all three sensors. The classification accuracies are firstly tested for normality using the Shapiro–Wilk test to determine suitable statistics e.g. parametric or non-parametric statistics.

4. Experimentations and results

4.1. Data collection

In order to test the proposed method, we collected the sensor data from a group of elderly people. The data collection was carried out in a living-home in Chiang Mai, Thailand in order to replicate natural living environment. This process had been carried out over several different days. Before the data collection process, the participants were given brief introduction about the study and an explanation about the data collection processes and written informed consents were obtained from all participants and they were informed they could withdraw at any time from the study. The project was approved by the Faculty of Computing, Engineering and Technology Academic Ethics Team, Staffordshire University, UK. The participants were asked about their personal health issues and evaluated using the Barthel Index (Mahoney & Barthel, 1965) to assess if they were suitable for participations.

A total of 19.2 h of sensor data were collected from 12 healthy older adults aged between 65 and 78 years. Characteristics of the participants including age, weight, height and Body Mass Index (BMI) are shown in Table 4. The participants were asked to wear two eZ430-Chronos watches on their wrists and perform 11 activities as listed in Table 1. The watch on the dominant wrist was set to transmit acceleration data while the other watch recorded temperature and altitude. The participant was asked to perform each activity for 5 min except for brushing teeth, dressing, walking downstairs and walking upstairs which had no time limit (see Fig. 5). The participant was allowed to perform these activities in any order and they can take breaks between activities. Before the data collection, the watches had been calibrated and paired with the PC. The researcher marked down the start, stop time and name of each activity. In order to reduce the strain caused by the appearance of the researcher during the data collection process, the participants were left to perform activities at their own paces without direct supervision. The acceleration data was collected using software developed on MatLab®. Temperature and altitude data were recorded on the watches internal memory which was later transferred to PC using the provided software from Texas Instruments (2010). The data collected from accelerometer were date, time, acceleration on X, Y and Z axis. The data collected from temperature sensor and altimeter were date, time, temperature and altitude. All collected data were later transferred to MatLab® for further transformation and analysis.

4.2. Data pre-processing

The sensor data was pre-processed using WMA and segmented at 3.88 s with 50% data overlapping resulting in a total of 17,843 patterns. It is noted that the number of walking upstairs and walking downstairs classes are relatively low. This was due to the fact that there were no time limit on these activities (on average, the participant used 5 s to climb up the 6-step stairs) in order to reflect natural behavior. Also, it was difficult to obtain large data set from these activities due to physical restriction because of their ages. The data from walking upstairs and downstairs classes only constitute 2% of all dataset which is clearly imbalanced. This will affect classification performance where most techniques assume samples are distributed evenly among different classes. Also, an imbalanced data set poses other problems such as difficulty in establishing accurate decision boundary, error in interpreting classification results, and data from minority class tend to be treated as noise (Nguyen, Bouzerdoum, & Phung, 2009).

In this work, it was decided to remove data from walking downstairs and walking upstairs classes as the numbers of samples were too low to be able to discover true classes boundaries especially in

Table 4
Participants characteristics.

Gender	Age (year)			Weight (kg)		Height (m)		BMI (kg/m^2)	
	Mean	Std.	Range	Mean	Std.	Mean	Std.	Mean	Std.
Female	72.11	4.54	13.00	48.26	10.13	1.53	0.060	20.53	4.10
Male	71.00	3.61	7.00	51.80	12.51	1.64	0.070	19.18	4.14
All	71.83	4.20	13.00	49.14	10.28	1.56	0.079	20.19	3.96

Fig. 5. Participants carry out activities naturally.

our case which showed highly overlapped classes. Also, according to the interview with the participants, it was found that majority of them live in a bungalow or on ground floor while participants who live on 2-floored houses only use stairs couple times a day (to access their bedrooms). The under-sampling technique was used to obtain a new dataset with the balanced number of samples from each class. All data from the smallest class i.e. dressing class were preserved. The same data size was obtained from the other eight classes. In total, the new balanced dataset contained $805 \times 9 = 7245$ patterns. This study used 10-fold cross validation which the data were randomly divided into 10 folds, one of 10 folds is used as the validation data, one of 10 folds is used as the test data and the remaining 8 folds are used as the training data in turn, the mean of the classification rates by using these 10 test data sets is used as the final classification rate.

Features are then generated from the raw sensor data as described in Section 3.3.3. Next, a feature selection is carried out using neural network and the proposed feature combination techniques in order to select the optimum feature subset. The neural network technique with resilient backpropagation and 20 hidden neurons was used. The feature ranking procedure using 10-fold cross validation and 10 runs was carried out. Fig. 6 shows that using the proposed feature combination method can achieved higher accuracy.The final subset was obtained by observing the truncation point of the mean accuracy of all datasets. It was decided to use 16 features as listed in Table 5.

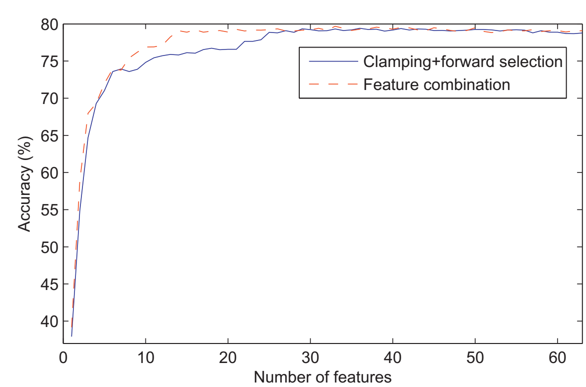

Fig. 6. Classification accuracy between using feature combination and clamping feature selection techniques.

4.3. Classification results

The neural network used in this experiment test was developed using MatLab Neural Network Toolbox®. The network had one hidden layer and the numbers of hidden nodes were selected based on the minimum error on validation sets. The RBF network used were built using MatLab Neural Network Toolbox®. The RBF parameters, SPREAD, which defines the radius of the RBF neurons were determined from 10-fold cross validation.

The classification results from MLP and RBF were not very good. The highest accuracy achieved by MLP was 81.52% while RBF only achieved 72.18%. SVM, on the other hand, showed statistically better classification performance. The analysis of confusion matrix of the three neural networks found that the classifiers made similar misclassifications thus combining three classifiers would not increase classification accuracy. Therefore, it was decided to use only SVM in the proposed method.

Table 5
The selected features used in this study.

Data source	Selected features		
Accelerometer	RMS *Y*-axis	RMS *X*-axis	Maximum *Y*-axis
	Minimum *Y*-axis	Difference *Z*-axis	Maximum *Z*-axis
	Key Coefficient *Y*-axis	Correlation X, Y	Minimum *Z*-axis
	Minimum *X*-axis	Maximum norm acc.	Difference *Y*-axis
Temperature sensor	Mean temperature	Key coefficient temperature	Min temperature
Altimeter	Entropy altitude		

SVM was applied which was constructed using LIBSVM (Chang & Lin, 2011) which is a free library for constructing the SVM model. A radial Gaussian kernel functions was used. For SVM parameters, a grid search (Hsu et al., 2003) using $C = 2^0, 2^{0.25}, \ldots, 2^{7.75}, 2^8$ and $\gamma = 2^1, 2^{1.25}, \ldots, 2^{3.75}, 2^4$ on validation sets were carried out using 10-fold cross validation with 10 runs. All combinations of C and γ were tested on each data set. The optimum parameters were selected based on the highest mean accuracy which were $C = 2^{2.5}$ and $\gamma = 2^3$.

A test using 100 unseen data sets were carried out. The proposed model achieved mean classification accuracy of 90.23%, standard deviation of 1.179, and standard error mean of 0.1179. When observing classification results of each class, the model also achieved high accuracy between 83.05% and 95.45%. Table 6 shows mean precision, recall and F-score of the nine classes. In general, the results showed high precision and recall indicating that the model was high performance (Precision = 90.43% ± 6.37%, Recall = 90.23% ± 5.06%). The average F-score of the proposed model was 0.9026 and standard deviation was 0.0567. The model did extremely well in detecting sleeping activity. Activities such as watching TV, sweeping, walking and feeding also had been detected very well. However, the model did not do well in detecting dressing activity.

Within the 9.77% of mean misclassification, the errors were mostly from dressing (19.27%), ironing (16.47%) and brushing teeth (16.41%) and washing dishes (15.91%) classes. Table 7 shows the confusion matrix of the proposed method. The numbers with the underlines show results from the model that achieved the lowest

Table 6
Mean precision and recall of the proposed method.

	Brush teeth	Dress/undress	Feed	Iron	Sleep	Sweep	Walk	Wash dishes	Watch TV
Precision	0.8495	0.7907	0.9041	0.8554	0.9814	0.9573	0.9494	0.8904	0.9606
Recall	0.8556	0.8305	0.9304	0.8551	0.9545	0.9538	0.9334	0.8600	0.9470
F-score	0.8517	0.8091	**0.9166**	0.8546	**0.9676**	**0.9553**	**0.9409**	0.8741	**0.9534**

Table 7
Confusion matrix of the proposed method.

Actual	Predicted								
	Brush teeth	Dress/undress	Feed	Iron	Sleep	Sweep	Walk	Wash dishes	Watch TV
Brush teeth	$\underline{64}$ 68.45 $\overline{73}$	$\underline{3}$ 3.03 $\overline{1}$	$\underline{5}$ 2.57 $\overline{1}$	$\underline{4}$ 2.44 $\overline{1}$	$\underline{1}$ 0.52 $\overline{0}$	$\underline{0}$ 0 $\overline{0}$	$\underline{0}$ 0.02 $\overline{0}$	$\underline{1}$ 2.37 $\overline{2}$	$\underline{2}$ 0.060 $\overline{2}$
Dress/undress	$\underline{6}$ 3.23 $\overline{4}$	$\underline{57}$ 66.44 $\overline{68}$	$\underline{4}$ 1.23 $\overline{1}$	$\underline{4}$ 3.31 $\overline{1}$	$\underline{0}$ 0.07 $\overline{0}$	$\underline{3}$ 1.63 $\overline{1}$	$\underline{3}$ 1.97 $\overline{2}$	$\underline{2}$ 1.85 $\overline{3}$	$\underline{1}$ 0.27 $\overline{0}$
Feed	$\underline{4}$ 1.96 $\overline{1}$	$\underline{2}$ 0.76 $\overline{2}$	$\underline{72}$ 74.43 $\overline{75}$	$\underline{0}$ 0.74 $\overline{0}$	$\underline{0}$ 0.27 $\overline{0}$	$\underline{0}$ 0.02 $\overline{0}$	$\underline{0}$ 0.08 $\overline{0}$	$\underline{1}$ 0.80 $\overline{0}$	$\underline{1}$ 0.94 $\overline{2}$
Iron	$\underline{5}$ 2.36 $\overline{2}$	$\underline{4}$ 4.92 $\overline{2}$	$\underline{1}$ 0.77 $\overline{1}$	$\underline{65}$ 68.41 $\overline{72}$	$\underline{0}$ 0.10 $\overline{0}$	$\underline{1}$ 0.18 $\overline{0}$	$\underline{0}$ 0.14 $\overline{1}$	$\underline{3}$ 2.70 $\overline{2}$	$\underline{1}$ 0.42 $\overline{0}$
Sleep	$\underline{0}$ 0.72 $\overline{0}$	$\underline{2}$ 1.13 $\overline{1}$	$\underline{1}$ 0.71 $\overline{0}$	$\underline{4}$ 0.29 $\overline{0}$	$\underline{72}$ 76.36 $\overline{79}$	$\underline{0}$ 0.09 $\overline{0}$	$\underline{0}$ 0.18 $\overline{0}$	$\underline{0}$ 0.25 $\overline{0}$	$\underline{1}$ 0.27 $\overline{0}$
Sweep	$\underline{0}$ 0.03 $\overline{0}$	$\underline{0}$ 1.49 $\overline{1}$	$\underline{0}$ 0.11 $\overline{0}$	$\underline{0}$ 0.10 $\overline{1}$	$\underline{0}$ 0.08 $\overline{0}$	$\underline{77}$ 76.30 $\overline{76}$	$\underline{2}$ 1.57 $\overline{1}$	$\underline{1}$ 0.23 $\overline{0}$	$\underline{0}$ 0.09 $\overline{1}$
Walk	$\underline{1}$ 0.24 $\overline{0}$	$\underline{1}$ 3.18 $\overline{2}$	$\underline{0}$ 0.11 $\overline{0}$	$\underline{0}$ 0.14 $\overline{0}$	$\underline{0}$ 0.04 $\overline{0}$	$\underline{3}$ 1.34 $\overline{0}$	$\underline{75}$ 74.67 $\overline{78}$	$\underline{0}$ 0.09 $\overline{0}$	$\underline{0}$ 0.19 $\overline{0}$
Wash dishes	$\underline{4}$ 2.72 $\overline{3}$	$\underline{2}$ 2.82 $\overline{1}$	$\underline{2}$ 1.27 $\overline{1}$	$\underline{1}$ 3.85 $\overline{4}$	$\underline{0}$ 0.12 $\overline{0}$	$\underline{0}$ 0.04 $\overline{0}$	$\underline{0}$ 0.01 $\overline{0}$	$\underline{71}$ 68.80 $\overline{71}$	$\underline{0}$ 0.37 $\overline{0}$
Watch TV	$\underline{0}$ 1.00 $\overline{1}$	$\underline{1}$ 0.46 $\overline{0}$	$\underline{2}$ 1.22 $\overline{2}$	$\underline{1}$ 0.80 $\overline{0}$	$\underline{1}$ 0.26 $\overline{0}$	$\underline{1}$ 0.14 $\overline{0}$	$\underline{0}$ 0.07 $\overline{0}$	$\underline{1}$ 0.29 $\overline{0}$	$\underline{73}$ 75.76 $\overline{77}$

Note: $\underline{n}$ indicates minimum and $\bar{n}$ indicates maximum.

accuracy, the number with the bars show results from the highest accuracy, and the mean values are in between those two numbers.

The confusion matrix revealed that the model often confused dressing class with ironing (24.41%) or brushing teeth (23.82%) classes. Ironing activities were also frequently misclassified as dressing (42.45%), washing dishes (23.30%) or brushing teeth (20.36%) activities. Classification of brushing teeth was regularly confused with dressing (26.23%), feeding (22.25%), ironing (21.13%) and washing dishes (20.52%).

4.4. Hypotheses testing result

4.4.1. H1: The proposed method can achieve more than 90% accuracy

The result from the Shapiro–Wilk test indicates that the data were normally distributed (SW = 0.979, df = 99, p = 0.107). The result of the null hypothesis testing that $H_0 : \mu \leqslant 90.00$ indicated that the accuracy difference was significant at the 5% level on a one-tailed test (T = 2.336, df = 99, p = 0.0296). Therefore, the null hypothesis is rejected in favor of the experimental hypothesis that the mean accuracy of the proposed method is higher than 90% indicating that the method can accurately detect elderly ADLs.

4.4.2. H2: Classification accuracy can be improved by combining data from temperature sensor and/or altimeter with accelerometer

In order to control the experiment, the number of input was set to 16 to comply with the number of selected features used in our proposed method. For configuration A, 16 top accelerometer features based on feature rankings in Section 3.3.4 were selected. We selected the best combination of features for both configurations B and C through the experimentations. 16 features in configuration D was selected from the proposed method (see Table 5).

Classifications were conducted using SVM on 10-fold cross validation $\times$ 10 times = 100 data sets. Optimum SVM parameters C and γ were selected using the grid search for each configuration. The results of classification accuracy using features from only accelerometer, accelerometer with temperature sensor, and accelerometer with altimeter, and combination of these sensors are shown in Table 8. The result showed that classification accuracy was increased when temperature or altimeter is combined with accelerometer. The data were tested for normality using the Shapiro–Wilk test which revealed that the data were not normally distributed ($p < 0.001$). Thus, it was appropriate to use non-parametric statistics for hypothesis testing. The Kruskal–Wallis test at 5% significance level was used to test the null hypothesis that the median classification accuracies are the same across all configurations.The result indicated that there was a statistically significant difference in median of accuracies between different configurations ($H(3) = 305.730$, $p < 0.001$) with a mean rank of 50.59 for using only accelerometer, 170.49 for using accelerometer with temperature sensor, 265.06 for using accelerometer with altimeter, and 315.87 for using combination of all three sensors.

A further pair-wise comparison between configuration D and others e.g. A VS D, B VS D, C VS D were conducted using the Mann–Whitney U test. The comparison results indicated that there was a statistically significant difference in median accuracy between configuration A and D ($U = 0.00$, $p < 0.001$), B and D ($U = 660.00$, $p < 0.001$), C and D ($U = 2803.50$, $p < 0.001$). The results also indicated that the mean rank of configuration D was significantly higher than other configurations. Therefore, it can be concluded that by combining data from temperature sensor and/or altimeter with accelerometer, classification accuracy can be improved. The result showed that using a combination of accelerometer, temperature sensor and altimeter achieved the highest classification accuracy among other configurations.

5. Discussion and conclusions

The main aim of this study is to develop an activity recognition method for detection of nine daily activities of an elderly person by considering on both technical and practical aspects. Taking the practical aspect into account, the proposed recognition method uses small, inexpensive, non-intrusive non-stigmatize multiple sensors i.e. accelerometer, temperature sensor and altimeter embedded on a wrist watch as inputs. Sensors embedded on an everyday object i.e. wrist watch would diminish the stigmatization which is one of the important concerns in ageing-support equipment. Also, wrist-worn equipment reduces issues such as restriction or discomfort that may happen when performing activities suitable for long-term activity recognition in free-space living environment. The selected sensors are small and low-cost which make them ideal for a non-intrusive and cost-effective assisted living solution.

As for the technical aspect, it is also important that the proposed activity detection method is able to achieve high performance. Different classification models were compared based on MLP, RBF and SVM. The results indicated that SVM is the most powerful classification algorithm. Therefore, we proposed a wrist-worn multi-sensors based activity recognition and classification method for detecting elderly ADLs using SVM. The proposed method achieved high classification performance of F-score between 0.81 and 0.97 and overall accuracy of 90.23%. This demonstrated that the proposed method performs very well on detecting activities of an elderly person. The method can detect several daily activities including BADLs such as feeding, brushing teeth, dressing, walking sleeping and IADLs such as washing dishes, ironing, sweeping floor and watching TV.

The confusion matrix revealed that the proposed method often gets confused among dressing, ironing, brushing teeth and washing dishes activities. Dressing is the most difficult activity to be detected as there is no clear pattern on how this activity should be performed e.g. one participant could undress/dress her top first while the other could do the opposite. Finding a generalized decision boundary for the dressing activity proved to be challenging. For the other three classes i.e. ironing, brushing teeth and washing dishes, the results could be implied that these classes have some common characteristics. Ironing, brushing teeth and washing

Table 8
Classification accuracies of using different set of sensors.

Configuration	Sensor	Selected features	$\overline{Accuracy}$(%)
A	Accelerometer	RMS_Y, RMS_X, MAX_Y, MIN_Y, MIN_Z, MIN_X, DIF_Y, MAX_Z, KEY_Y, COR_{XY}, MAX_{NORM}, DIF_Y, $MEAN_Y$, MAX_X, $MEAN_Z$, $RANGE_Z$	82.7694
B	Accelerometer, Temperature	RMS_Y, RMS_X, MAX_Y, MIN_Y, MIN_Z, MIN_X, DIF_Y, MAX_Z, KEY_Y, $MEAN_{TEMP}$, STD_{TEMP}, MAX_{TEMP}, MIN_{TEMP}, $RANGE_{TEMP}$, ENT_{TEMP}, KEY_{TEMP}	87.5764
C	Accelerometer, Altimeter	RMS_Y, RMS_X, MAX_Y, MIN_Y, MIN_Z, MIN_X, DIF_Y, MAX_Z, KEY_Y, $MEAN_{ALT}$, STD_{ALT}, MAX_{ALT}, MIN_{ALT}, ENE_{ALT}, ENT_{ALT}, KEY_{ALT}	89.3736
D	Accelerometer, Temperature, Altimeter	RMS_Y, RMS_X, MAX_Y, MIN_Y, MIN_Z, MIN_X, DIF_Y, MAX_Z, KEY_Y, COR_{XY}, MAX_{NORM}, DIF_Y, ENT_{ALT}, $MEAN_{TEMP}$, MIN_{TEMP}, KEY_{TEMP}	90.2250

dishes are all involved in some kind of repetitive stroke motion e.g. back-and-forth motion. The analysis showed that maximum and minimum acceleration on Y-axis of ironing and washing dishes activities are highly overlapped. Nevertheless, when comparing to previous works (Fleury et al., 2010), our method still achieved a higher classification result on these activities. Also, the proposed method can accurately detect less active activities i.e. sleeping and watching TV (F-score = 95.45% and 96.04%, respectively comparing to 93.9%).

The results from the experiment indicated that, in our application, accelerometer is the most valuable sensor for activity recognition. This result supports the previous finding from the literature (Györbíró et al., 2009). The temperature sensor and altimeter when using on their own did not achieve good classification results compared to the accelerometer. However, the experimental results revealed that by adding information from temperature sensor or altimeter, the classification performance is statistically improved, and that the combination of all three sensors resulted in the highest classification accuracy confirming our hypothesis. This result supports the theory that a variable that is completely useless by itself can provide significant performance improvement when taken with others (Guyon & Elisseeff, 2003). Features from accelerometer when taken with features from temperature and altitude improve class separability resulting in better classification performances.

Our method from both technical and practical aspects demonstrates high potential to be successfully applied in assisted living systems in several aspects. Firstly, the method only uses wearable sensors comparing to other works which are involved in the installation of specialized systems or sensors in house or on objects such as works by (Fleury et al., 2010; Hong et al., 2010, 2009; Huang et al., 2010). The installation or retrofitting of dedicated equipment would lead to the higher costs of assisted living solutions made it less likely to be affordable by the general population. Our choices of sensors are inexpensive and the watch only cost £37 (Texas Instruments, 2010) which would make it a cost-effective solution. Lower sampling rate (33 Hz) was also used to reduce power consumption thus extending battery life which helps to lower the number of time to change battery. Our approach also does not require a large number of sensors e.g. RFID tags to be attached to objects for easy set up.

Secondly, small sensors used in this study were embedded in a wrist watch which diminishes stigmatization and increases user acceptance. In other systems such as Hong et al. (2010), Huang et al. (2010) where users are required to wear a special glove which may be perceived as stigma are more difficult to be accepted by elderly people. User acceptance is one of the key issues in assisted living systems. Unless it is accepted by users, even high performance systems would consider failures. A wrist watch is an everyday object which can simply be part of normal dressing accessories.

Thirdly, non-intrusive set of sensors were used to allow unobtrusive activity monitoring. Our choice of sensors are non-visual and also presented on a single location i.e. wrist reduces the chance of it becoming an obstacle or interfering with daily activities.

Finally, accuracy is the most important technical aspect of any detection system. It is also one of the key concerns when adopting technology in assisted living systems. The results of the study showed that our proposed multi-sensor worn on wrist can be used to detect Activities of Daily Livings with high accuracy $(\overline{Accuracy} = 90.23\%, \overline{F-score} = 0.9026)$. Table 9 shows a comparison between the proposed method and previous works. The proposed method can achieve comparable or even higher accuracy comparing to previous works considering the sensor locations and number of recognised activities.

The proposed method can detect several activities occurred in a typical daily life which can be applied in various elderly care applications. Fig. 7 illustrates examples of applications of the activity recognition for elderly care. For example, it can be used in an activity monitoring system to automatically monitor an elderly person who lives independently. This system could be integrated with other communication-related systems e.g. web-based application allowing a series of detected activities to be observed by families or close friends. This would help ensuring the elderly safety and well-being, providing peace of mind. This system is particularly useful for elderly people who lives away from their families. This work can also benefit nursing home or any care institutes. As the number of elderly people increases, the ratio of caregivers per patient could be affected. The proposed ADLs detection can be used to provide real-time monitoring so that care can be done in a more effective and efficient way.

The detection of elderly ADLs can be used to produce activity records automatically for long-term observation. This method allows unobtrusive continuous activity recording in a cost-effective way,

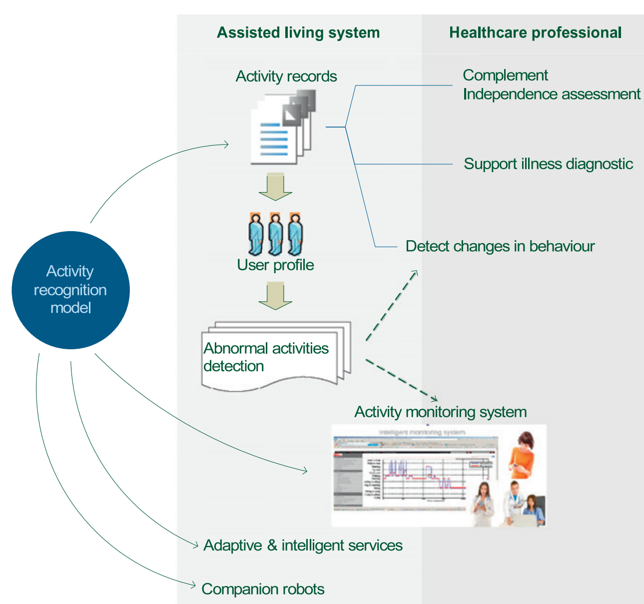

Fig. 7. Application of activity recognition in elderly care.

Table 9
Accuracy between proposed method and previous works.

Author	Recognized activities	Brush teeth	Dress/undress	Feed	Sleep	Walk	Wash dish	Iron	Average
Proposed model	9	85.56%	83.05%	93.04%	95.45%	93.34%	86.00%	85.51%	90.23%
Fleury et al. (2010)	7	64.30%[a]	75.00%	97.80%	93.9%	95.00%	–	–	86.20%
Maurer et al. (2006)	6	–	–	–	–	>90%	–	–	87.10%
Huang et al. (2010)	4	85.00%	–	84.00%	–	–	76.00%	–	82.00%
Hong et al. (2010)	–	–	–	–	92.66%	84.36%	–	97.94%	–

[a] Activities include wash hand and teeth were detected.

while it is usually expensive for direct observation. Long-term activity records can help detect changes in activity pattern or behavior which could be used to identify decline in health. Early detection in behavior changes can help prevent or minimize risk in long term health damage. The activity record can also be used as additional information for illness diagnostic by health care professionals. In addition, further analysis of long-term activity record can be used to produce normal activity pattern for elderly individuals allowing abnormal activities to be detected. The elderly would have better opportunity to live safely and independently which is particularly important for people who suffer from severe illnesses e.g. Parkinsons or Alzheimers (Yin, Yang, & Pan, 2008). Detection of unusual activities can also be used as a first indicator when the elderly develop cognitive decline or symptoms of illness or even injury.

The proposed method can be used as an alternative or complementary of an assessment of elderly people independence level. The level of independence of an elderly person is usually judged by physicians interviewing patients using some types of paper-based form e.g. Barthel Index. The analysis of the record of detected activities can be used to complement the interview information and to confirm the assessment. It can be used to provide level of independence assessment automatically over time.

The proposed method can be a part of assisted living systems providing more intelligent and adaptive solutions. For example, it can be used to help the elderly recognize and cope with the cognitive decline associated with illness and aging by sending adaptive personalized activity reminders (Pollack et al., 2003). The systems will help older persons adapt to cognitive decline and continue the satisfactory performance of their routine activities and potentially enabling them to remain in their own homes longer. Also, it can be used in a companion robot for elderly living alone in their homes. The proposed work can be used to determine users activity in order that the robot can respond in an appropriate way.

In conclusion, the study presented an activity recognition method which considers both technical and practical aspects focusing on user acceptance, privacy (non-visual), systems accuracy and cost. These factors are vital keys toward intelligent assisted living solution for elderly people who are the most important stakeholders. In the era of ageing population, intelligent assisted living could help elderly people maintain living independently in their own homes. The proposed method can detect nine ADLs of an elderly person using multi-sensor worn on wrist. The results of the study demonstrate the method is generally high performance and practical. In future work, the investigation will be carried out in order to improve the method performance, particularly on those activities which often misclassified by adding other sensors. Further study should be carried out to assess the validity of the proposed method in free-living environment.

Acknowledgments

The authors thank Mrs. Pan Ayumak and Mrs. Juntip Tumtibsri, the representatives of Watket elderly club, Chiang Mai, Thailand for supporting the data collection process. We also would like to thank all participants who participated in the data collection and Mr. Somchit Chernbumroong who provided location for the study. This work has been supported by the EU Erasmus Mundus Project-eLINK (EM ECW-Ref.149674-EM-1-2008-1-UK-ERAMUNDUS).

References

Aarts, E. (2004). Ambient intelligence: A multimedia perspective. *IEEE Multimedia, 11*, 12–19. ID: 1.

Banos, O., Damas, M., Pomares, H., Prieto, A., & Rojas, I. (2012). Daily living activity recognition based on statistical feature quality group selection. *Expert Systems with Applications, 39*, 8013–8021.

Bayer, A., Harper, L., Greenwald, M., & of Retired Persons A. A. (2000). Fixing to stay: A national survey of housing and home modification issues. American Association of Retired Persons.

Botia, J. A., Villa, A., & Palma, J. (2012). Ambient assisted living system for in-home monitoring of healthy independent elders. *Expert Systems with Applications, 39*, 8136–8148.

Chaaraoui, A. A., Climent-Prez, P., & Flrez-Revuelta, F. (2012). A review on vision techniques applied to human behaviour analysis for ambient-assisted living. *Expert Systems with Applications, 39*, 10873–10888.

Chang, C.-C., & Lin, C.-J. (2011). Libsvm: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology, 2*, 27:1–27:27.

Chernbumroong, S., Atkins, A., & Yu, H. (2011). Activity classification using a single wrist-worn accelerometer. In *Fifth international conference on software, knowledge information, industrial management and applications (SKIMA), 2011* (pp. 1–6). IEEE.

Chernbumroong, S., Atkins, A. S., & Yu, H. (2010). Perception of smart home technologies to assist elderly people. In *The fourth international conference on software, knowledge, information management and applications (SKIMA 2010)* (pp. 90–97).

Costa, A., Castillo, J. C., Novais, P., Fernndez-Caballero, A., & Simoes, R. (2012). Sensor-driven agenda for intelligent home care of the elderly. *Expert Systems with Applications, 39*, 12192–12204.

Demiris, G., Hensel, B. K., Skubic, M., & Rantz, M. (2008). Senior residents' perceived need of and preferences for "smart home" sensor technologies. *International Journal of Technology Assessment in Health Care, 24*, 120–124.

Demiris, G., Rantz, M., Aud, M., Marek, K., Tyrer, H., Skubic, M., et al. (2004). Older adults' attitudes towards and perceptions of'smart home'technologies: A pilot study. *Informatics for Health and Social Care, 29*, 87–94.

Fleury, A., Vacher, M., & Noury, N. (2010). Svm-based multimodal classification of activities of daily living in health smart homes: Sensors, algorithms, and first experimental results. *IEEE Transactions on Information Technology in Biomedicine, 14*, 274–283. ID: 1.

Fuentes, D., Gonzalez-Abril, L., Angulo, C., & Ortega, J. (2012). Online motion recognition using an accelerometer in a mobile device. *Expert Systems with Applications, 39*, 2461–2465.

Guyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection. *The Journal of Machine Learning Research, 3*, 1157–1182.

Györbı́ró, N., Fábián, A., & Hományi, G. (2009). An activity recognition system for mobile phones. *Mobile Networks and Applications, 14*, 82–91.

Hong, X., Nugent, C., Mulvenna, M., McClean, S., Scotney, B., & Devlin, S. (2009). Evidential fusion of sensor data for activity recognition in smart homes. *Pervasive and Mobile Computing, 5*, 236–252.

Hong, Y. J., Kim, I. J., Ahn, S. C., & Kim, H. G. (2010). Mobile health monitoring system based on activity recognition using accelerometer. *Simulation Modelling Practice and Theory, 18*, 446–455.

Hsu, C.-W., Chang, C.-C., & Lin, C.-J. (2003). A practical guide to support vector classification. Technical report. Department of Computer Science, National Taiwan University.

Hsu, C. W., & Lin, C. J. (2002). A comparison of methods for multiclass support vector machines. *IEEE Transactions on Neural Networks, 13*, 415–425.

Huang, P.-C., Lee, S.-S., Kuo, Y.-H., & Lee, K.-R. (2010). A flexible sequence alignment approach on pattern mining and matching for human activity recognition. *Expert Systems with Applications, 37*, 298–306.

Kleinberger, T., Becker, M., Ras, E., Holzinger, A., & Muller, P. (2007). Ambient intelligence in assisted living: enable elderly people to handle future interfaces. In C. Stephanidis (Ed.), *Universal access in human-computer interaction: Ambient interaction UAHCI'07* (pp. 103–112). Berlin, Heidelberg: Springer-Verlag.

Kwon, O., Shim, J. M., & Lim, G. (2012). Single activity sensor-based ensemble analysis for health monitoring of solitary elderly people. *Expert Systems with Applications, 39*, 5774–5783.

Mahoney, F., & Barthel, D. (1965). Functional evaluation: The barthel index. *Maryland State Medical Journal, 14*, 56–61.

Martı́nez-Pérez, F., González-Fraga, J., Cuevas-Tello, J., & Rodrı́guez, M. (2012). Activity inference for ambient intelligence through handling artifacts in a healthcare environment. *Sensors, 12*, 1072–1099.

Maurer, U., Rowe, A., Smailagic, A., & Siewiorek, D. (2006). Location and activity recognition using ewatch: A wearable sensor platform. *Ambient Intelligence in Everyday Life*, 86–102.

Najafi, B., Aminian, K., Paraschiv-Ionescu, A., Loew, F., Bula, C. J., & Robert, P. (2003). Ambulatory system for human motion analysis using a kinematic sensor: Monitoring of daily physical activity in the elderly. *IEEE Transactions on Biomedical Engineering, 50*, 711–723. ID: 1.

Nguyen, H. G., Bouzerdoum, A., & Phung, S. L. (2009). Learning pattern classification tasks with imbalanced data sets. *Pattern Recognition*, 193–208.

Olivieri, D. N., Conde, I. G., & Sobrino, X. A. V. (2012). Eigenspace-based fall detection and activity recognition from motion templates and machine learning. *Expert Systems with Applications, 39*, 5935–5945.

Parkka, J., Ermes, M., Korpipaa, P., Mantyjarvi, J., Peltola, J., & Korhonen, I. (2006). Activity classification using realistic data from wearable sensors. *IEEE Transactions on Information Technology in Biomedicine, 10*, 119–128.

Poh, M. (2011). Continuous assessment of epileptic seizures with wrist-worn biosensors. Ph.D. thesis. Massachusetts Institute of Technology.

Pollack, M. E., Brown, L., Colbry, D., McCarthy, C. E., Orosz, C., Peintner, B., et al. (2003). Autominder: An intelligent cognitive orthotic system for people with memory impairment. *Robotics and Autonomous Systems, 44*, 273–282.

Texas Instruments, T.I. (2010). eZ430-Chronos development tool user's guide. Manual SLAU292C.

United Nations, D. o. E., & Social Affairs, P. D. (2011). World population prospects: The 2010 revision, cd-rom edition.

Wang, W., Jones, P., & Partridge, D. (2000). Assessing the impact of input features in a feedforward neural network. *Neural Computing and Applications, 9*, 101–112.

Yin, J., Yang, Q., & Pan, J. J. (2008). Sensor-based abnormal human-activity detection. *IEEE Transactions on Knowledge and Data Engineering, 20*, 1082–1090.

Yu, H., Spenko, M., & Dubowsky, S. (2003). An adaptive shared control system for an intelligent mobility aid for the elderly. *Autonomous Robots, 15*, 53–66.

Zhu, C., & Sheng, W. (2011). Motion- and location-based online human daily activity recognition. *Pervasive Mobile Computing, 7*, 256–269.