Applications and Challenges of Human Activity Recognition using Sensors in a Smart Environment

Jubil T Sunny  
Department of Computer Science and Engineering  
St. Joseph’s College of Engineering and Technology  
Palai, Kerala, India  

Sonia Mary George  
Department of Computer Science and Engineering  
St. Joseph’s College of Engineering and Technology  
Palai, Kerala, India  

Jubilant J Kizhakkethottam  
Department of Computer Science and Engineering  
St. Joseph’s College of Engineering and Technology  
Palai, Kerala, India  

Abstract

We are currently using smart phone sensors to detect physical activities. The sensors which are currently being used are accelerometer, gyroscope, barometer, etc. Recently, smart phones, equipped with a rich set of sensors, are explored as alternative platforms for human activity recognition. Automatic recognition of physical activities – commonly referred to as human activity recognition (HAR) – has emerged as a key research area in human-computer interaction (HCI) and mobile and ubiquitous computing. One goal of activity recognition is to provide information on a user’s behavior that allows computing systems to proactively assist users with their tasks. Human activity recognition requires running classification algorithms, originating from statistical machine learning techniques. Mostly, supervised or semi-supervised learning techniques are utilized and such techniques rely on labeled data, i.e., associated with a specific class or activity. In most of the cases, the user is required to label the activities and this, in turn, increases the burden on the user. Hence, user-independent training and activity recognition are required to foster the use of human activity recognition systems where the system can use the training data from other users in classifying the activities of a new subject.

Keywords: Human Activity Recognition

I. INTRODUCTION

Mobile phones or smart phones are rapidly becoming the central computer and communication device in people’s lives. Smart phones, equipped with a rich set of sensors, are explored as an alternative platform for human activity recognition in the ubiquitous computing domain. Today’s Smartphone not only serves as the key computing and communication mobile device of choice, but it also comes with a rich set of embedded sensors [1], such as an accelerometer, digital compass, gyroscope, GPS, microphone, and camera. Collectively, these sensors are enabling new applications across a wide variety of domains, such as healthcare, social networks, safety, environmental monitoring, and transportation, and give rise to a new area of research called mobile phone sensing. Human activity recognition systems using different sensing modalities, such as cameras or wearable inertial sensors, have been an active field of research. Besides the inclusion of sensors, such as accelerometer, compass, gyroscope, proximity, light, GPS, microphone, camera, the ubiquity, and unobtrusiveness of the phones and the availability of different wireless interfaces, such as WI-Fi, 3G and Bluetooth, make them an attractive platform for human activity recognition. The current research in activity monitoring and reasoning has mainly targeted elderly people, or sportsmen and patients with chronic conditions.

The percentage of elderly people in today’s societies keep on growing. As a consequence, the problem of supporting older adults in loss of cognitive autonomy who wish to continue living independently in their home as opposed to being forced to live in a hospital. Smart environments have been developed in order to provide support to the elderly people or people with risk factors who wish to continue living independently in their homes, as opposed to live in an institutional care. In order to be a smart environment, the house should be able to detect what the occupant is doing in terms of one’s daily activities. It should also be able to detect possible emergency situations. Furthermore, once such a system is completed and fully operational, it should be able to detect anomalies or deviations in the occupant’s routine, which could indicate a decline in his abilities. In order to obtain accurate results, as much information as possible must be retrieved from the environment, enabling the system to locate and track the supervised person in each moment, to detect the position of the limbs and the objects the person interacts or has the intention to interact with. Sometimes, details like gaze direction or hand gestures [1] can provide important information in the process of analyzing the human activity. Thus, the supervised person must be located in a smart environment, equipped with devices such as sensors, multiple view cameras or speakers.
Although smart phone devices are powerful tools, they are still passive communication enablers rather than active assistive devices from the user’s point of view. The next step is to introduce intelligence into these platforms to allow them to proactively assist users in their everyday activities. One method of accomplishing this is by integrating situational awareness and context recognition into these devices. Smart phones represent an attractive platform for activity recognition, providing built-in sensors and powerful processing units. They are capable of detecting complex everyday activities of the user (i.e., Standing, walking, biking) or the device (i.e., Calling), and they are able to exchange information with other devices and systems using a large variety of data communication channels.

Mobile phone sensing is still in its infancy. There is little or no consensus on the sensing architecture for the phone. Common methods for collecting and sharing data need to be developed. Mobile phones cannot be overloaded with continuous sensing commitments that undermine the performance of the phone (e.g., by depleting battery power). It is not clear what architectural components [4] should run on the phone. Individual mobile phones collect raw sensor data from sensors embedded in the phone. Information is extracted from the sensor data by applying machine learning and data mining techniques. These operations occur either directly on the phone. Where these components run could be governed by various architectural considerations, such as privacy, providing user real-time feedback, reducing communication cost between the phone and cloud, available computing resources, and sensor fusion requirements.

The rest of the paper is organized as follows: Section II presents some existing methods. Section III describes important sensors used for human activity recognition. Chapter IV represents various challenges and applications of activity recognition. Conclusions are presented in Chapter V.

II. RELATED WORKS

Activity recognition became an important research issue related to the successful realization of intelligent pervasive environments. It is the process by which an actor’s behavior and his or her environment are monitored and analyzed to infer the activities. Activity recognition consists of activity modeling, behavior and environment monitoring, data processing and pattern recognition [6]. Activity recognition systems typically have three main components:

- A low-level sensing module that continuously gathers relevant information about activities using microphones, accelerometers, light sensors, and so on
- A feature processing and selection module that processes the raw sensor data into features that help discriminate between activities
- A classification module that uses the features to infer what activity an individual or group of individuals is engaged
- In, for example, walking, cooking, or having a conversation.

There are several approaches for activity recognition as described as follows.

A. Vision-Based Activity Recognition

It uses visual sensing facilities: camera-based surveillance systems to monitor an actor’s behavior and the changes in its environment. It is composed of four steps: human detection, behavior tracking, activity recognition and high-level activity evaluation. Various other research approaches used different methods such as: single camera or stereo and infrared to capture activity context. This image-based approaches use single or multiple cameras to reconstruct the 3D human pose [5], to detect the coordinates of the joints and to extract the limbs of the body. The image analysis is possible by isolating the human body from the background. This is achieved using a the background subtraction algorithm that adapts to the environmental changes.

B. Sensor-Based Activity Recognition

It uses sensor network technologies to monitor an actor’s behavior along with its environment. In this case there are sensors attached to humans. Data from the sensors are collected and analyzed using data mining or machine learning algorithms to build activity models and perform activity recognition. In this case, they’re recognized activities included human physical movements: walking, running, sitting down/up as in. Most of wearable sensors are not very suitable for real applications due to their size or battery life. In sensor-based approach, can use either wearable sensors or object-attached sensors. The most used machine learning is the Hidden Markov Model (HMM)[19] – a graphical oriented method to characterize real world observations in terms of state models. Another good alternative is the Conditional Random Field (CRF) model, which is an un-directed graphical method which allow the dependencies between observations and the use of incomplete information about the probability distribution of a certain observable.

C. Human-Sensing Taxonomy

Classify under the large umbrella of “human-sensing” the process of extracting any information regarding the people in any environment. This describes the inference of spatiotemporal properties (STPs) only. These consist of low-level components regarding the position and history of people in an environment. More specifically:

1) Presence: Is there at least one person present? Presence is arguably the property that is most commonly sought- after in existing real-world applications, the most popular presence-sensor being motion sensors (PIR) and proximity
sensors (scalar infrared rangefinders). In cooperative scenarios, though, where people can be instrumented with portable or wearable devices, solutions such as RFID (radio-frequency identification) are becoming increasingly common.

2) Count: How many people are present? The number of people in an environment can be inferred by either employing a person-counting sensor (or sensors) that cover the entire area of interest, or by counting people at all the entry and exit points. Commercial people-counting solutions range from thermal imagers [SenSource] and break-beams, to simple mechanical barriers such as turnstiles.

3) Location: Where is each person? Location-detection, or "localization", consists of obtaining the spatial coordinates of a person's Centre of mass. Localization can be achieved using instrumented (such as GPS) or fully un-instrumented solutions (such as cameras). In addition, since a grid of presence sensors can also be used to localize people, localization can be considered a higher-resolution generalization of presence detection.

4) Track: Where was this person before? Tracking is the process of solving the correspondence problem, that is, extracting the spatio-temporal history of each person in a scene. Equivalently, tracking may be described as recovering a person's relative identity. For example, if upon detection a person is labeled with a temporary ID (e.g. 'person 6") then tracking is the problem of specifying at each subsequent sampling of the scene which detection is the same 'person 6". This temporary ID is typically lost in the presence of sensing gaps, such as when the person leaves the scene and returns on the next day. At that point, yesterday's 'person 6" will be given a new ID when re-detected. Situations that lead to the loss of a person's relative ID are often called ambiguities. In the remainder of this text, it will use the term piecewise tracking to qualify a tracker that is not capable of adequately handling ambiguities.

5) Identity: Who is each person? Is this person John? At a first glance it may seem odd to group 'identity" into the category of spatio-temporal properties. However, identification is nothing more than a natural extension of tracking where each person is always assigned the same globally unique ID rather than solely relative IDs. Therefore, identity-detection extends tracking so that it becomes possible to recover a person's spatial temporal history even across sensing gaps.

### III. SENSORS

Modern mobile phones come with a variety of sensors that automate or easy many of our daily tasks. In activity recognition use a different type of sensors as the source of the raw data collection. Sensors are mainly of three categories: video sensors [3], environmental-based sensors [7], and wearable sensors. Video sensors are basically cameras that are installed in the fixed places such as the entrance/exit of the Public places (to detect people’s appearance and actions), or in the living rooms or bedrooms (to track the users’ daily life). Cameras are also embedded in robots for a more active visual data capture. Visual monitoring for activity recognition is used in many applications such as surveillance, anti-terrorists, and anti-crime securities as well as life logging and assistance. Environment-based sensors [2] are used to detect the users’ interaction with the environment.

They are radio based sensors like WiFi, Bluetooth, and the infrared sensors. These sensors are usually deployed in indoor places such as office buildings or homes. They passively monitor the appearance of users at a certain location, or the users’ interaction with objects that are also equipped with sensors. Their limitations are that (1) they can only be applied to certain fixed locations, and (2) the cost for the full deployment of such sensors is often very high. Wearable sensors are the mobile sensors that are in small size and designed to be worn on the human body in daily activities. They can record users’ physiological states such as location changes, moving directions, speed, etc. Such sensors include accelerometers, microphones, GPS, barometers, etc. Most of the mobile sensors are equipped with smart phones.

#### A. Accelerometer

Accelerometers in mobile phones are used to detect the orientation or to sense the acceleration event of Smart phones. An accelerometer measures linear acceleration of movement. The reading includes three axes whose directions are pre-defined as in Fig 3.1.1. The acceleration is the raw data stream from the accelerometer. A set of vectors represents the raw data as: $\mathbf{Acc}_i = < x_i, y_i, z_i >$, $i = 1,2,3,...$. An accelerometer will measure the directional movement of a device but will not be able to resolve its lateral orientation or tilt during that movement accurately.

![Accelerometer axes on smart phones](image)

*Fig. 3.1.1: Accelerometer axes on smart phones*
The three axes readings are combined with a time stamp. In order to sample the frequency, most of existing accelerometers provide a user interface so that the user could choose a better sampling rate through experiments. Single and multi-axis accelerometers detect the combined magnitude and direction of linear, rotational and gravitational acceleration [8]. They can be used to provide limited motion sensing functionality. For example, a device with an accelerometer can detect rotation from vertical to horizontal in a fixed state location. As a result, accelerometers are primarily used for simple motion sensing applications in consumer devices such as changing the screen of a mobile device from portrait to landscape orientation. Its popularity is due to the fact that it directly measures the subject’s physiology motion status. For example, if a user changes his/her activity from walking to jogging, it will reflect on the signal shape of the accelerated reading along the vertical axis [9] there will be an abrupt change in the amplitude. Moreover, the acceleration data could indicate the motion pattern within a given time period, which is helpful in the complex activity recognition.

B. Compass Sensor
Magnetic sensors commonly referred to as compasses detect magnetic fields and measure their absolute position relative to the Earth’s magnetic north and nearby magnetic materials. Information from magnetic sensors can also be used to correct errors from other sensors such as accelerometers. It provides mobile phones with a simple orientation in relation to the Earth’s magnetic field. As a result, your phone always knows which way is North so it can auto-rotating your digital maps depending on your physical orientation. One example of how compass sensors are used in consumer devices is reorienting a displayed map to match up with the general direction a user is facing. Compass is a traditional tool to detect the direction with respect to the north-south pole of the earth by the use of magnetism. The compass sensor for smart phones works with a similar functionality. Figure 3.2.1 shows the compass reading display screen on a smart phone.

The raw data reading from a compass sensor is the float number between 0 degree and 360 degree. It begins from 0 degrees at the absolute north and the actual reading indicates the angle between current smart phone heading direction and the absolute north in clockwise. For example, the reading of heading to absolute East is 90 degrees and heading to absolute West is 270 degrees. The data stream returned from compass sensors is a set of floating numbers indicating the angle, campy (i=0,1,2,3,... 0 degree <= campy <= 360 degrees). Compass reading can be used to detect the direction change in the user’s motion such as walking.

C. Gyroscope
Gyroscopes measure the angular rate of rotational movement about one or more axes. Gyroscopes can measure complex motion accurately in multiple dimensions, tracking the position and rotation of a moving object, unlike accelerometers, which can only detect the fact that an object has moved or is moving in a particular direction. Further, unlike accelerometers and compasses, gyroscopes are not affected by errors related to external environmental factors such as gravity and magnetic fields. Hence, gyroscopes greatly enhance the motion sensing capabilities of devices and are used for advanced motion sensing applications in consumer devices such as full gesture and movement detection and simulation in video gaming. Gyroscope measures the phone’s rotation rate by detecting the roll [10], pitch, and yaw motions of the Smart phones along the x, y, and z axis, respectively. The axes directions are shown in Fig. 3.3.1.

The raw data stream from a gyroscope sensor is the rate of the rotation in rad/s (radian per second) around each of the three physical axes: Rotationi = <xi ; yi ; zi>(i=1,2,3...). The gyroscope is helpful in the navigation applications as well as some smart phone games which use the rotation data. Inactivity recognition research, the gyroscope is used to assist the mobile orientation detection.
D. Barometer
Pressure Sensors, also known as barometers measure relative and absolute altitude through the analysis of changing atmospheric pressure. Pressure sensors[15] can be used in consumer devices for sports and fitness or location-based applications where information on elevation can be valuable. The barometer is one of the latest sensors equipped on some advanced smart phones (e.g., Samsung Galaxy S4 and Google Nexus 4/10). It measures the atmospheric pressure of the environment that the sensor is placed in.

The air pressure varies with different altitude or Fig. 3.4.1 The barometer is one of the latest sensors equipped on some advanced smart phones (e.g., Samsung Galaxy S4 and Google Nexus 4/10). It measures the atmospheric pressure of the environment that the sensor is placed in. The air pressure varies with different altitude or Fig. 3.4.1 Even with places of the same altitude, but having different structures (e.g., narrow and wide hallways) inside a building. Thus, barometer reading can be used to indicate the user’s position change in localization related activity recognition.

IV. CHALLENGES AND APPLICATIONS

Although the research on activity recognition is beneficial from the mobile sensors’ unobtrusiveness, flexibility, and many other advances, it also faces challenges that have brought with them. In this section, review the major, common challenges for activity recognition using mobile sensors, and the corresponding solutions to alleviate them in the current literature.

A. Subject sensitivity
The accuracy of activity recognition, especially those based on the accelerometer data, is heavily affected by the subjects participated in training and testing stages. This is mainly due to the fact that different people have different motion patterns. Even for the same subject, she/he may have different patterns at different time. The comparative experiments show that training and testing on the same subject achieves the highest accuracy. Training and testing on the same group of multiple subjects has the second highest accuracy. The accuracy decreases when the test data are collected from same subject, but on different days. The
lowest accuracy is in the setting where the training data is collected from one subject on one day and testing is conducted on another subject on a different day. A recognition model trained on such a diversified dataset works more reliably when it is tested on data from new individuals. Deng et al.[16] proposed a cross-person activity recognition model to eliminate the effect of user sensitivity. The model training stage consists of two parts: The initial model is trained off-line and the adaptive model is updated online. For new users in the online phase, the algorithm selects those high confident recognition results in order to generate the new training data set. Based on this new training data set, the algorithm will update the recognition model to alleviate the subject sensitivity.

B. Location Sensitivity
Due to the property of accelerometer both in wearable sensors and smart phones, its raw reading heavily depends on the sensors’ orientation and positions on the subject’s body. For example, when a user is walking while holding a phone in his/her hand, the moving data reading is quite different from the data reading if the phone is in his/her pocket. One solution is proposed in Ref. [17] to address the orientation sensitivity by using another sensor: magnetometer. The magnetic field sensor provides the magnetic vector along three axes of the device’s coordinate system in the orthogonal directions. Hence, it could be utilized to derive the devices’ azimuth angle. Then the accelerometer reading can be converted to the earth coordinating axes reading. Park et al.[11] presented a device pose classification method based on the regularized kernel algorithm. It provides a way of how to estimate the smart phone’s pose before doing any motion data analysis.

C. Activity Complexity
The complexity of user activities also brings an additional challenge to the recognition model. For example, the motion during the transition period between two activities is difficult for the underlying classification algorithm to recognize. People performing multiple tasks[13] at the same time might also confuse the classifier which is trained under one activity-per-segment assumption. In addition, culture and individual difference might result in the variation in the way that people perform tasks, which in turn brings the difficulty in applying the activity recognition models globally. HMM is a natural solution to address the activity complexity by “smoothing” the error during the activity transition period.

D. Energy and Resource Constrains
Activity recognition applications require continuous sensing as well as online updating of the classification model, both of which are energy consuming. For the online is updating, it might also require significant computing resources (e.g., mobile phone memories). Based on the observation that the required sampling frequency differs for different activities, A3R[14] adaptively makes the choices on both sampling frequency and classification features. In this way, it reduces both energy and computing resource cost. It also removes the time-consuming frequency-domain feature calculation.

E. Insufficient Training Set
As mentioned in the subject sensitivity challenge part, it is highly desirable that the training data must contain as many varieties of the subjects as possible. However, it is not easy to coordinate people of different ages and body shapes to collect data under a controlled lab environment, not to mention the varieties of the environment itself. Semi- supervised learning is applied to address this issue. In many classification tasks, the unlabeled data, when used in conjunction with a small amount of labels data, can produce considerable improvement in learning accuracy. For activity recognition, the collection of unlabeled data is easy and requires near zero users’ effort. By combining semi- supervised learning with virtual evidence boosting (EVB) method, it reduces the human labelling cost as well as improves the efficiency for feature selection. Besides the traditional semi-supervised learning method, the scale-invariant classifier with “R” metric (SIC-R). SIC-R is designed to recognize multi scale events of human activities. The introduced feature descriptor of time-scale invariance allows the feature from one training set to describe events of the same semantics class which may take place over varying time scales. In this way, it reduces the demand on the training set.

Activity recognition is a core building block behind many interesting applications. The applications of mobile activity recognition can be classified according to their targeted beneficial subjects:

1) Applicable for the end users such as fitness tracking, health monitoring, fall detection, behavior-based context-awareness, home and work automation, and self-managing system;
2) Applications for the third parties such as targeted advertising, research platforms for the data collection, corporate management[18], and accounting;
3) Applications for the crowds and groups such as social networking and activity-based crowdsourcing. In this section, review some representative applications.

F. Daily Life Monitoring
Applications in daily life monitoring usually aim to provide a convenient reference for the activity logging, or assisting with exercise and healthy lifestyles. These devices are equipped with the embedded sensors such as accelerometer, gyroscope, GPS; and they track people’s steps taken, stairs climbed, calorie burned, hours slept, distance travelled, quality of sleep, etc. An online service is provided for users to review data tracking and visualization in reports. Compared with smart phone sensors, these
devices are more sophisticated, since their sensors are designed specifically for the activity detection and monitor. The drawback is that they are much more expensive. Smartphone applications with activity recognition techniques have been shown up in recent years as an alternative solution. These applications usually have similar roles as above specialized devices. They track users’ motion logs such as jogging route, steps taken, and sleeping time. By mining the logged data, they may offer the user a summary on his/her lifestyle and report the sleeping quality.

G. Personal Biometric Signature
A subject’s motion pattern is usually exclusive and unique. For example, when people raise their hands, it is almost impossible for two people’s hands to share the exact same motion patterns. Even in a successful imitation, the differences still exist because of the difference in the motion related bones and muscles on human bodies. Sensors such as accelerometers can capture those differences. The activity recognition techniques provide a possible solution for human biometric signature with patterns in motion/gestures. In these applications, pattern recognition methods are used to obtain the unique motion patterns, which are in turn saved in the database. It is convenient and feasible because of the pervasive usage of mobile devices. On the other side, the motion signature could also be used in a malicious way. For example, people could use the learned patterns to crack users’ behaviors, such as smart phone keyboard typing, or other spying activities.

H. Elderly and Youth Care
There is a growing need in elderly care (both physically and mentally), partially because of the retirement of the baby Boomer generation. A major goal of the current research in human activity monitoring is to develop new technologies and applications for elderly care. Those applications could help prevent harm, e.g., to detect older people’s dangerous situations. An architecture on the smart phone is developed with the purpose of users’ fall detection. Activity recognition and monitor sensors could help elders in a proactive way, such as life routine reminder (e.g., taking medicine), living activity monitoring for a remote robotic assists. The youth care is another field that benefits from the activity recognition research. Applications include monitoring infants’ sleeping status and predicting their demands for food or other stuff. Activity recognition techniques are also used in children’s (ASD) detection.

I. Localization
Activity recognition on mobile phones could help with context-awareness and hence can be applied in localization. One reason to use mobile sensors rather than GPS for localization is that GPS signal is usually very weak inside buildings and underground. On the other hand, the activity recognition techniques with mobile sensors could assist in locating the position. In addition, GPS localization is 2-D based positioning which has no information about a user’s altitude. Activity recognition techniques for mobile phones could fill in this gap. A similar system is for infrastructure-free floor localization. A third reason to use mobile sensors for localization is that GPS accuracy decreases inside cities with tall buildings surrounded. In this situation, GPS-based localization might confuse between a movie theatre and a restaurant, which might be just too close to each other of the distance. Activity recognition related applications can alleviate this kind of mistakes by augmenting the positions with people’s current activity type.

J. Industry Manufacturing, Assisting
The activity recognition techniques could also assist workers in their daily work. This work (wearIT@Work) introduces wearable sensors into work — “wearable computing” is a kind of extension of the body that allows a worker to perform “extraordinary tasks”. Other applications based on activity recognition include smart cameras that can understand people’s gestures in the film shooting field, robot assistance in car production, etc.

V. CONCLUSION
Smart phones are ubiquitous and becoming more and more sophisticated. This has been changing the landscape of people’s daily life and has opened the doors for many interesting data mining applications. Human activity recognition is a core building block behind these applications. It takes the raw sensor’s reading as inputs and predicts a user’s motion activity. This paper presents a comprehensive survey of the recent advances in activity recognition with smart phone’s sensors. Here introduce the basic concepts of activity recognition (such as sensors, activity types, etc.).

Here review the core data mining techniques behind the mainstream activity recognition algorithms, analyze their major challenges, and introduce a variety of real applications enabled by activity recognition. The activity recognition based on smart phone sensors leads to many possible future research directions. Besides the applications mentioned in Section V, an even novel way could be equipping smart phones with intelligent applications to replace the traditional devices such as remote control, traffic controlling, and tracking devices. Smartphone applications that can recognize users’ gestures could send a corresponding command to home electronics. Thus, instead of keeping different remotes in one’s cabinet, can just install one application that has the remote functions. The cross field research could be developed in many fields because of the mobile activity recognition techniques.
REFERENCES


