Chapter 2
Sensor-Based Activity Recognition Review

2.1 Introduction

The idea of using sensors for activity monitoring and recognition has been existent since the late 90s. It was initially pioneered and experimented by the work of the Neural Network house [1] in the context of home automation, and a number of location-based applications aiming to adapt systems to users' whereabouts [2, 3]. The approach was soon found to be more useful and suitable in the area of ubiquitous and mobile computing—an emerging area in the late 90s, due to its easy deployment. As such, extensive research has been undertaken to investigate the use of sensors in various application scenarios of ubiquitous and mobile computing, leading to considerable work on context-awareness [4–6], smart appliances [40, 41] and activity recognition [7–10]. Most research at that time made use of wearable sensors, either dedicated sensors attached to human bodies or portable devices like mobile phones, with application to ubiquitous computing scenarios such as providing context-aware mobile devices. Activities being monitored in these researches are mainly physical activities like motion, walking and running. These early works lay a solid foundation for wearable computing and still inspire and influence today's research.

In the early 2000s, a new sensor-based approach that uses sensors attached to objects to monitor human activities appeared. This approach, which was later dubbed as the "dense sensing" approach, performs activity recognition through the inference of user-object interactions [11, 12]. The approach is particular suitable for dealing with activities that involve a number of objects within an environment, or instrumental Activities of Daily Living [13, 14]. Research on this approach has been heavily driven by the intensive research interests and huge research effort on smart home-based assistive living, such as the EU's AAL program [15]. In particular, sensor-based activity recognition can better address sensitive issues in assistive living such as privacy, ethics and obtrusiveness than conventional vision-based approaches. This combination of application needs and technological advantages has stimulated considerable research activities in a global scale, which gave rise to a large number of research projects, including the House_n, CASAS, Gator-Tech, inHaus, AwareHome,

DOMUS and iDorm projects, to name but a few. As a result of the wave of intensive investigation, there have seen a plethora of impressive works on sensor-based activity recognition in the past several years.

While substantial research has been undertaken, and significant progress has been made, the two main approaches, wearable sensors based and dense sensing based activity recognition are currently still focuses of study. The former is mainly driven by the ever-popular pervasive and mobile computing while the latter is predominantly driven by smart environment applications such as ambient assisted living. Interests in various novel applications are still increasing and application domains are rapidly expanding.

2.2 Sensor-Based Activity Monitoring

Currently a wide range of sensors, including contact sensors, RFID, accelerometers, audio and motion detectors, to name but a few, are available for activity monitoring. These sensors are different in types, purposes, output signals, underpinning theoretical principles and technical infrastructure. However, they can be classified into two main categories in terms of the way they are deployed in activity monitoring applications. These are wearable sensors and dense sensors, and are described in details in the following.

2.2.1 Wearable Sensor Based Activity Monitoring

Wearable sensors generally refer to sensors that are positioned directly or indirectly on a human body. They generate signals when the user performs activities. As a result, they can monitor features that are descriptive of the person's physiological state or movement. Wearable sensors can be embedded into clothes, eyeglasses, belts, shoes, wristwatches, mobile devices or positioned directly on the body. They can be used to collect information such as body position and movement, pulse, and skin temperature. Researchers have found that different types of sensor information are effective for classifying different types of activities. In the following, we summarise the common practice in wearable sensor-based activity monitoring.

Accelerometer sensors are probably the most frequently used wearable sensor for activity monitoring. They are particularly effective in monitoring actions that involve repetitive body motions, such as walking, running, sitting, standing, climbing stairs. Bao et al. [11] provide a summary of research work that recognises human activities using acceleration data. Kern et al. [16] deploy a network of 3-axis accelerometers distributed over the user's body. Each accelerometer provides information about the orientation and movement of the corresponding body part. Lukowicz et al. [17] recognize workshop activities using body-worn microphones and accelerometers. Measuring acceleration and angular velocity (the angle of the user's thigh) through

wearable sensors, such as accelerometers and gyroscopes, Lee et al. [10] propose a dead-reckoning method for determining a user's location and recognizing sitting, standing and walking behaviours. Mantyjarvi [18] recognises human ambulation and posture on acceleration data collected from the hip.

GPS sensors are another widely used wearable sensor for monitoring location-based activities in open pervasive and mobile environments. Patterson et al. [12] present details of detecting human high-level behaviour from a GPS sensor stream, such as boarding a bus at a particular bus stop, travelling and disembarking. Ashbrook et al. [19] use GPS to learn significant locations and predict movement across multiple users. Liao et al. [20] learn and infer a user's mode of transportation and their goal in addition to abnormal behaviours (e.g., taking a wrong bus) based on GPS data logs.

Biosensors are an emerging technology aiming to monitor activities through vital signs. A diversity of sensors in different forms has been studied in order to measure the wide range of vital signs such as blood pressure, heart rate, EEG, ECG and respiratory information. Sung et al. [21] monitor the body temperature of soldiers to detect hypothermia. Harms et al. [22] use information gathered by a smart garment to identify body posture.

In addition to the investigation of different wearable sensors for activity monitoring, research on the support and novel application of wearable computing has been undertaken. Pantelopoulos et al. [23] present a survey on wearable systems for monitoring and early diagnosis for the elderly. Dakopoulos and Bourbakis [24] present a survey on wearable obstacle avoidance electronic travel aids for visually impaired. Yoo et al. [25] design on-body and near-body networks that use the human body itself as a channel for creating BodyNets. Cooper and Au use wearable sensors to design and evaluate assistive wheelchairs [26] and smart walking sticks [27]. Kim et al. [28] use wearable sensors to recognize gestures. Madan et al. [29] characterize a person's social context by evaluating a user's proximity, speech, head movements and galvanic skin response.

Wearable sensor-based activity monitoring suffers from limitations. Most wearable sensors need to run continuously and be operated hands-free. This may have difficulties in real-world application scenarios. Practical issues include the acceptability or willingness to use wearable sensors and the viability and ability to wear them. Technical issues include the size, ease of use, battery life and effectiveness of the approach in real-world scenarios. To address these issues, vigorous investigation of smart garments has been carried out, which aims to embed sensors in garments for monitoring [30]. Another research thread is to make use of existing gadgets that have already been carried in a daily basis like smartphones as intelligent sensors for activity monitoring, recognition and assistance. This practice has been in place for a while and is expected to gain large-scale uptake given the latest development and affordability of such palm-held electronic devices.

Obviously, wearable sensors are not suitable for monitoring activities that involve complex physical motions and/or multiple interactions with the environment. In some cases, sensor observations from wearable sensors alone are not sufficient to differentiate activities involving simple physical movements (e.g., making tea and

making coffee). As a result, dense sensing based activity monitoring has emerged, which is described below.

2.2.2 Ambient Sensor Based Activity Monitoring

Ambient sensor based activity monitoring refers to the practice that a large number of ambient sensors are attached to objects within an environment, and activities are monitored by detecting user-object interactions. The approach is based on real-world observations that activities are characterized by the objects that are manipulated during their performance. A simple indication of an object being used can often provide powerful clues about the activity being undertaken. As such it is assumed that activities can be recognised from sensor data that monitors human interactions with objects in the environment. By dense sensing, we refer to the way and scale with which sensors are used. Using dense sensing a large number of sensors, normally low-cost low-power and miniaturized, are deployed in a range of objects or locations within an environment for the purpose of monitoring movement and behaviour.

As dense sensing-based monitoring embeds sensors within environments, this makes it more suitable for creating ambient intelligent applications such as smart environments. As such, dense sensing-based activity monitoring has been widely adopted in ambient assisted living (AAL), via the smart home paradigm [14]. Sensors in an SH can monitor an inhabitant's movements and environmental events so that assistive agents can infer the undergoing activities based on the sensor observations, thus providing just-in-time context-aware ADL assistance. For instance, a switch sensor in the bed can strongly suggest sleeping, and pressure mat sensors can be used for tracking the movement and position of people within the environment.

Since the introduction of the idea in the early 2000s, extensive research has been undertaken to investigate the applicability of the approach in terms of sensor types, modalities and applications. For example, Tapia et al. [64] use environmental state-change sensors to collect information about interaction with objects and recognize activities that are of interest to medical professionals such as toileting, bathing, and grooming. Wilson et al. [31] use four kinds of anonymous and binary sensors, motion detectors, break-beam sensors, pressure mats, and contact switches for simultaneous tracking and activity recognition. Wren et al. [32] employ networks of passive infrared motion sensors to detect the presence and movement of heat sources. With this captured data they can recognize low-level activities such as walking, loitering, and turning, as well as mid-level activities such as visiting and meeting. Srivastava et al. [33] exploit wireless sensor network to develop a smart learning environment for young children. Hollosi et al. [34] use voice detection techniques to perform acoustic event classification for monitoring in Smart Homes. Simple object sensors are adopted in [35].

Given the abundance of different types and modalities of sensors, sensors have been used in different ways and combinations for dense sensing activity monitoring in many application scenarios. It is impossible to claim that one sensor deployment

for a specific application scenario is superior to the other. The suitability and performance is usually down to the nature of the type of activities being assessed and the characteristics of the concrete applications. As such, in this chapter, we shall not discuss in detail the different usage of dense sensing in various scenarios but simply introduce its rationale as described above.

Generally speaking, wearable sensor-based activity monitoring receives more attention in mobile computing while dense sensing is more suitable for intelligent environment enabled applications. It is worth pointing out that wearable sensors and dense sensing are not mutually exclusive. In some applications, they have to work together. For example, RFID (Radio Frequency Identification) based activity monitoring requires that objects are instrumented with tags and users wear an RFID reader affixed to a glove or a bracelet. Philipose and Fishkin [36, 37] developed two devices, iGlove and iBracelet, working as wearable RFID readers that detect when users interact with unobtrusively tagged objects. Patterson et al. [38] performed fine-grained activity recognition (i.e., not just recognising that a person is cooking but determining what they are cooking) by aggregating abstract object usage. Hodges et al. [39] proposed to identify individuals from their behaviour based on their interaction with the objects they use in performing daily activities. Buettner et al. [40] recognize indoor daily activities by using an RFID sensor network. In most cases, wearable sensors and dense sensing are complementary and can be used in combination in order to yield optimal recognition results. For example, Gu et al. [41] combine wearable sensors and object sensors for collecting multimodal sensor information. Through a pattern-based method, they recognize sequential, interleaved and concurrent activities.

While substantial research has been undertaken, and significant progress has been made, the two main approaches, wearable sensors based and dense sensing-based activity recognition are currently still focuses of study. The former is mainly driven by the ever-popular pervasive and mobile computing while the latter is predominantly driven by smart environment applications such as ambient assisted living. Interests in various novel applications are still increasing and application domains are rapidly expanding.

2.3 Data-Driven Approaches to Activity Modelling and Recognition

Data-driven activity modeling can be classified into two main categories: generative and discriminative. In the generative approach, one attempts to build a complete description of the input or data space, usually with a probabilistic model such as a Bayesian network. In the discriminative approach, one only models the mapping from inputs (data) to outputs (activity labels). Discriminative approaches include many heuristic (rule-based) approaches, neural networks, conditional random fields

and linear or non-linear discriminative learning (e.g. support vector machines). In the following, we cover major results using each of these methods.

2.3.1 Generative Methods

The simplest possible generative approach is the naïve Bayes classifier, which has been used with promising results for activity recognition. Naïve Bayes classifiers model all observations (e.g. sensor readings) as arising from a common causal source: the activity, as given by a discrete label. The dependence of observations on activity labels is modelled as a probabilistic function that can be used to identify the most likely activity given a set of observations. Despite the fact that these classifiers assume conditional independence of the features, the classifiers yield good accuracy when large amounts of sample data are provided. Nevertheless, naïve Bayes classifiers do not explicitly model any temporal information, usually considered important in activity recognition.

The Hidden Markov Model (HMM) is probably the most popular generative approach that includes temporal information. A HMM is a probabilistic model with a particular structure that makes it easy to learn from data, to interpret the data once a model is learned, and is both easy and efficient to implement. It consists of a set of hidden (latent) states coupled in a stochastic Markov chain, such that the distribution over states at some time depends only on the values of states at a finite number of preceding times. The hidden states then probabilistically generate observations through a stochastic process. HMMs made their impact initially through use in the speech recognition literature, where latent states correspond to phoneme labels, and observations are features extracted from audio data. HMMs have more recently been adopted as a model of choice in computer vision for modelling sequential (video) data. HMM use a Markov chain over a discrete set of states. A closely relative of the HMM uses continuous states, a model usually referred to as a linear dynamical system (LDS). State estimation in LDSs is better known as a Kalman filter. LDSs have been used with inputs from a variety of sensors for physiological condition monitoring [42] in which a method is also introduced to deal with unmodelled variations in data, one of the major shortcomings of the generative approach.

HMMs form the basis of statistical temporal models. They are, in fact, a special case of the more general dynamic Bayesian networks (DBNs), which are Bayesian networks in which a discrete time index is explicitly represented. Inference and learning in DBNs is simply an application of network propagation in Bayesian networks. DBNs usually make a Markovian assumption, but explicitly represent conditional independencies in the variables, allowing for more efficient and accurate inference and learning. A well-known early use of DBNs for activity monitoring was in the Lumière project, where a Microsoft Windows user's need for assistance was modelled based on their activities on the screen [43].

A simple DBN extension of HMMs is the coupled HMM for recognition of simultaneous human actions. Coupled Hidden Markov Models (CHMMs) have two Marko-

vian chains, each modelling a different stream of data, with a coupling between them to model their inter-dependence. Oliver et al. [57] learn a multi-layer model of office activity to choose actions for a computational agent. The model uses multimodal inputs, making only very slight use of computer vision. The Assisted Cognition project [44] has made use of DBNs, in particular for Opportunity Knocks [20], a system designed to provide directional guidance to a user navigating through a city. This system uses a three level hierarchical Markov model represented as a DBN to infer a user's activities from GPS sensor readings. Movement patterns, based on the GPS localization signals, are translated into a probabilistic model using unsupervised learning. From the model and the user's current location, future destinations and the associated mode of transportation can be predicted. Based on the prediction, the system has the ability to prompt the user if an error in route is detected.

Wilson and Atkeson [31] use DBNs to simultaneously track persons and model their activities from a variety of simple sensors (motion detectors, pressure sensors, switches, etc.). DBNs were also used in the iSTRETCH system [45], a haptic robotic device to assist a person with stroke rehabilitation. The DBN models the person's current behaviours, their current abilities, and some aspects of their emotional state (e.g. their responsiveness, learning rate and fatigue level). The person's behaviours correspond to how long they take for each exercise, what type of control they exhibit and whether they compensate. These behaviours are inferred from sensors on the device and in the person's chair.

Even though they are simple and popular, HMMs and DBNs have some limitations. A HMM is incapable of capturing long- range or transitive dependencies of the observations due to its very strict independence assumptions (on the observations). Furthermore, without significant training, a HMM may not be able to recognize all of the possible observation sequences that can be consistent with a particular activity.

2.3.2   Discriminative Methods

A drawback of the generative approach is that enough data must be available to learn the complete probabilistic representations that are required. In this section, we discuss an alternative approach for modelling in which we focus directly on solving the classification problem, rather than on the representation problem. The complete data description of a generative model induces a classification boundary, which can be seen by considering every possible observation and applying the classification rule using inference. The boundary is thus implicit in a generative model, but a lot of work is necessary to describe all the data to obtain it. A discriminative approach, on the other hand, considers this boundary to be the primary objective.

Perhaps the simplest discriminative approach is Nearest Neighbor (NN), in which a novel sequence of observations is compared to a set of template sequences in a training set, and the most closely matching sequences in the training set vote for their activity labels. This simple approach can often provide very good results. Bao and Intille [11] investigated this method along with numerous other base-level classifiers

for the recognition of activities from accelerometer data. They found that the simple nearest neighbor approach is outperformed by decision trees, a related method, where the training data is partitioned into subsets according to activity labels and a set of rules based on features of the training data. The rules can then be used to identify the partition (and hence the activity label) corresponding to a new data sample. Maurer et al. [46], employed decision trees to learn logical descriptions of activities from complex sensor readings from a wearable device (the eWatch). The decision tree approach offers the advantage of generating rules that are understandable by the user, but it is often brittle when high precision numeric data is collected. Stikic and Schiele [47] use a clustering method in which activities are considered as a "bag of features" to learn template models of activities from data with only sparse labels.

Many discriminative approaches explicitly take into account the fact that, for classification, it is actually only the points closest to the boundary that are of interest. The ones very far away (the "easy" ones to classify) do not play such a significant role. The challenge is therefore to find these "hard" data points (the ones closest to the boundary). These data points will be known as the "support vectors", and actually define the boundary. A support vector machine (SVM) is a machine learning technique to find these support vectors automatically. A recent example of an SVM in use for activity modelling is presented by Brdiczka et al. [48] where a model of situations is learned automatically from data by first learning roles of various entities using SVMs and labelled training data, then using unsupervised clustering to build 'situations' or relations between entities, which are then labelled and further refined by end users. The key idea in this work is to use a cognitive model (situation model) based on cognitive theory motivated by models of human perception of behaviour in an environment. The CareMedia project [49] also uses an SVM to locate and recognize social interactions in a care facility from multiple sensors, including video and audio. The fusion of video and audio allowed 90% recall and 20% precision in identifying interactions including shaking hands, touching, pushing and kicking. The CareMedia project's goals are to monitor and report behaviour assessments in a care home to caregivers and medical professionals.

Ravi et al. also found that SVMs performed consistently well, but also investigated meta-level classifiers that combined the results of multiple base-level classifiers [50]. Features extracted from worn accelerometers are extracted and classified using five different base-level classifiers (decision tables, decision trees, k-nearest neighbors, SVM and Naïve Bayes). The meta-level classifiers are generated through a variety of techniques such as boosting, bagging, voting, cascading and stacking. For recognizing a set of eight activities including standing, walking, running, going up/down stairs, vacuuming and teeth brushing, they found that a simple voting scheme performed the best for three easier experimental settings, whereas boosted SVM performed best for the most difficult setting (test/training separation across users and days).

In practice, many activities may have non-deterministic natures, where some steps of the activities may be performed in any order, and so are concurrent or interwoven. A conditional random field (CRF) is a more flexible alternative to the HMM that addresses such practical requirements. It is a discriminative and generative probabilistic model that represents the dependence of a hidden variable y on an observed

variable *x*. Both HMMs and CRFs are used to find a sequence of hidden states based on observation sequences. Nevertheless, instead of finding a joint probability distribution $p(x,y)$ as the HMM does, a CRF attempts to find only the conditional probability $p(y|x)$. A CRF allows for arbitrary, non-independent relationships among the observation sequences, hence the added flexibility. Another major difference is the relaxation of the independence assumptions, in which the hidden state probabilities may depend on the past and even future observations. A CRF is modelled as an undirected acyclic graph, flexibly capturing any relation between an observation variable and a hidden state. CRFs are applied to the problem of activity recognition in [51] where they are compared to HMMs, but only in a simple simulated domain. Liao et al. [52] use hierarchical CRFs for modelling activities based on GPS data. Hu and Yang [53] use skip-chain CRFs, an extension in which multiple chains interact in a manner reminiscent of the CHMM, to model concurrent and interleaving goals, a challenging problem for activity recognition. Mahdaviani and Choudhury [54] show how semi-supervised CRFs can be used to learn activity models from wearable sensor data.

2.3.3   Heuristic and Other Methods

Many approaches do not fall clearly into discriminative or generative categories, but rather use a combination of both, along with some heuristic information. The Independent Lifestyle Assistant (ILSA) is an example, as it uses a combination of heuristic rules and statistical models of sequential patterns of sensor firings and time intervals to help a person with planning and scheduling [55]. PEAT (the Planning and Execution Assistant and Trainer) is a cognitive assistant that runs on a mobile device and helps compensate for executive functional impairment. PEAT uses reactive planning to adjust a user's schedule based on their current activities. Activity recognition in PEAT is based on what the user is doing, and on data from sensors on the mobile device. These are fed into an HMM, the outputs of which are combined with the reactive planning engine [56].

Other work has investigated how activities can be modelled with a combination of discriminative and generative approaches [57], how common sense models of everyday activities can be built automatically using data mining techniques [58], and how human activities can be analysed through the recognition of object use, rather than the recognition of human behaviour [59]. This latter work uses DBNs to model various activities around the home, and a variety of radio frequency identification (RFID) tags to bootstrap the learning process. Some authors have attempted to compare discriminative and generative models [11, 50], generally finding the discriminative models yield lower error rates on unseen data, but are less interpretable. Gu et al. [41] use the notion of emerging patterns to look for frequent sensor sequences that can be associated with each activity as an aid for recognition. Omar et al. [60] present a comparative study of a variety of classification methods for analysing multi-modal sensor data from a smart walker.

The generative approach, which attempts to build a complete description of the input or data space, usually with probabilistic analysis methods such as Markov models [61] and Bayesian networks [48] for activity modelling. These methods incorporate an inhabitant's preferences by tuning the initial values of the parameters of the probabilistic models. The major disadvantage with such methods is that the model is static and subjective in terms of probabilistic variable configuration. An alternative approach is referred to as the discriminative approach, which only models the mapping from inputs (data) to outputs (activity labels). Discriminative approaches include many heuristics (rule-based) approaches, for example, neural networks, linear or non-linear discriminant learning. They use machine learning techniques to extract ADL patterns from observed daily activities, and later use the patterns as predictive models [48]. Both approaches require large datasets for training models, thus suffer from the data scarcity or the "Cold Start" problem. It is also difficult to apply modelling and learning results from one person to another.

2.4 Knowledge-Driven Approaches to Activity Modelling and Recognition

Knowledge-driven activity recognition and modelling is motivated by real-world observations that for most activities of daily living and working, the list of objects required for a particular activity is limited and functionally similar. Even if the activity can be performed in different ways the number and type of these involved objects do not vary significantly. For example, it is common sense that the activity "*make coffee*" consists of a sequence of actions involving a *coffee pot, hot water, a cup, coffee, sugar* and *milk*; the activity "*brush teeth*" contains actions involving a *toothbrush, toothpaste, water tap, cup* and *towel*. On the other hand, as humans have different lifestyles, habits or abilities, they may perform various activities in different ways. For instance, one may like strong *white coffee,* and another may prefer a special brand of coffee. Even for the same type of activity (e.g., making *white coffee*), different individuals may use different items (e.g., *skimmed milk* or *whole milk*) and in different orders (e.g., adding *milk* first and then *sugar*, or vice versa). Such domain-dependent activity-specific prior knowledge provides valuable insights into how activities can be constructed in general and how they can be performed by individuals in specific situations.

Similarly, knowledge-driven activity recognition is founded upon the observations that most activities, in particular, routine activities of daily living and working, take place in a relatively specific circumstance of time, location and space. The space is usually populated with events and entities pertaining to the activities, forming a specific environment for specific purposes. For example, brushing teeth is normally undertaken twice a day in a bathroom in the morning and before going to bed and involves the use of toothpaste and a toothbrush; meals are made in a kitchen with a cooker roughly three times a day. The implicit relationships between activities,

related temporal and spatial context and the entities involved (objects and people) provide a diversity of hints and heuristics for inferring activities.

Knowledge-driven activity modelling and recognition intends to make use of rich domain knowledge and heuristics for activity modelling and pattern recognition. The rationale is to use various methods, in particular, knowledge engineering methodologies and techniques, to acquire domain knowledge. The captured knowledge can then be encoded in various reusable knowledge structures, including activity models for holding heuristics and prior knowledge in performing activities, and context models for holding relationships between activities, objects and temporal and spatial contexts. Comparing to data-driven activity modelling that learns models from large-scale datasets and recognises activities through data intensive processing methods, knowledge-driven activity modelling avoids a number of problems, including the requirement for large amounts of observation data, the inflexibility that arises when each activity model needs to be computationally learned, and the lack of reusability that results when one person's activity model is different from another's.

Knowledge structures can be modelled and represented in different forms, such as schemas, rules or networks. This will decide the way and the extent to which knowledge is used for following processing such as activity recognition, prediction and assistance. In terms of the manner in which domain knowledge is captured, represented and used, knowledge-driven approaches to activity modelling and recognition can be roughly classified into three main categories as presented in the following sections.

2.4.1   Mining-Based Approach

The rationale of a mining-based approach is to create activity models by mining existing activity knowledge from publicly available sources. More specifically, given a set of activities, the approach seeks to discover from the text corpuses a set of objects used for each activity and extract object usage information to derive their associated usage probabilities. The approach essentially views the activity model as a probabilistic translation between activity names (e.g., "*make coffee*") and the names of involved objects (e.g., "*mug*", "*milk*"). As the correlations between activities and their objects are common-sense prior knowledge (e.g., most of us know how to carry out daily activities), such domain knowledge can be gleaned in various sources such as how-tos (e.g., those at ehow.com), recipes (e.g., from epicurious.com), training manuals, experimental protocols, and facility/device user manuals.

A mining-based approach consists of a sequence of distinct tasks. Firstly, it needs to identify activities of concern and relevant sources that describe these activities. Secondly, it uses various methods, predominantly information retrieval and analysis techniques, to retrieve activity definitions from specific sources and extract phrases that describe the objects used during the performance of the activity. Then algorithms, predominantly probabilistic and statistical analysis methods such as co-occurrences and association are used to estimate the object-usage probabilities. Finally, the mined

object and usage information is used to create activity models such as a HMM that
can be used further for activity recognition.

Mining-based activity modelling was initially investigated by researchers from
Intel Research [62, 63]. Perkowitz et al. [63] proposed the idea of mining the Web for
large-scale activity modelling. They used the QTag tagger to tag each word in a sen-
tence with its part of speech (POS) and a customized regular expression extractor to
extract objects used in an activity. They then used the Google Conditional Probabil-
ities (GCP) APIs to determine automatically the probability values of object usage.
The mined object and their usage information are then used to construct DBN mod-
els through Sequential Monte Carlo (SMC) approximation. They mined the website
ehow.com for roughly 2300 directions on performing domestic tasks (from "*boiling
water in the microwave*" to "*change your air filter*"), and the website ffts.com and
epicurious.com for a further 400 and 18,600 recipes respectively, generating a total
21,300 activity models. Using the DBN activity models they have performed activ-
ity recognition for a combination of real user data and synthetic data. While initial
evaluation results were positive, the drawback was that there are no mechanisms to
guarantee the mined models capturing completely the sequence probabilities and the
idiosyncrasy of certain activities. The inability to capture such intrinsic characteris-
tics may limit the model's accuracy in real deployments.

Wyatt et al. [62] followed Perkowitz's approach by mining the Web to create DBN
activity models. However, this group extended the work in three aspects, aiming to
address the idiosyncrasies and to improve model accuracy. To cover the wide variety
of activity definition sources, they mined the Web in a more discriminative way in
a wider scope. They did this by building a specialized genre classifier trained and
tested with a large number of labelled Web pages. To enhance model applicability,
they used the mined models as base activity models and then exploited the Viterbi
Algorithm and Maximum Likelihood to learn customized activity parameters from
unsegmented, unlabelled sensor data. In a bid to improve activity recognition accu-
racy they also presented a bootstrap method that produced labelled segmentations
automatically. Then they used the Kullback–Leibler (KL) divergence to compute
activity similarity.

A difficulty in connecting mined activities with tagged objects is that the activity
models may refer to objects synonymously. For example, both a "*mug*" and "*cup*"
can be used for making *tea*; both a "*skillet*" and "*frying pan*" be used for making
pasta. This leads to a situation that one activity may have different models with
each having the same activity name but different object terms. To address this, Tapia
et al. [64] proposed to extract collections of synonymous words for the functionally-
similar objects automatically from WordNet, an online lexical reference system for
the English language. The set of terms for similar objects is structured and represented
in a hierarchical form known as the object ontology. With the similarity measure
provided by the ontology, an activity model will not only cover a fixed number of
object terms but also any other object terms that are in the same class in the ontology.

Another shortcoming of early work in the area is that the segmentation is carried
out in sequential order based on the duration of an activity. As the duration of per-
forming a specific activity may vary substantially from one to another, this may give

rise to applicability issues. In addition, in sequential segmentation, one error in one segment may affect the segmentations of the subsequent traces. To tackle this, Palmes et al. [65] proposed an alternate method for activity segmentation and recognition. Instead of relying on the order of object use, they exploited the discriminative trait of the usage frequency of objects in different activities. They constructed activity models by mining the Web and extracting relevant objects based on their weights. The weights are then utilized to recognize and segment an activity trace containing a sequence of objects used in a number of consecutive and non-interleaving activities. To do this, they proposed an activity recognition algorithm, *KeyExtract*, which uses the list of discriminatory key objects from all activities to identify the activities present in a trace. They further proposed two heuristic segmentation algorithms, *MaxGap* and *MaxGain*, to detect the boundary between each pair of activities identified by *KeyExtract*. Boundary detection is based on the calculation, aggregation, and comparison of the relative weights of all objects sandwiched in any two key objects representing adjacent activities in a trace. Though the mining-based approach has a number of challenges relating to information retrieval, relation identification and the disambiguation of term meaning, nevertheless, it provides a feasible alternative to model a large amount of activities. Initial research has demonstrated the approach is promising.

Mining-based approaches are similar to data-driven approaches in that they all adopt probabilistic or statistical activity modelling and recognition. But they are different from each other in the way the parameters of the activity models are decided. The mining-based approaches make use of publicly available data sources avoiding the "cold start" problem. Nevertheless, they are weak in dealing with the idiosyncrasies of activities. On other hand, data-driven approaches have the strength of generating personalized activity models, but they suffer from issues such as "cold start" and model reusability for different users.

2.4.2   Logic-Based Approach

The rationale of logical approaches is to exploit logical knowledge representation for activity and sensor data modelling, and to use logical reasoning to perform activity recognition. The general procedure of a logical approach includes (1) to use a logical formalism to explicitly define and describe a library of activity models for all possible activities in a domain, (2) to aggregate and transform sensor data into logical terms and formula, and (3) to perform logical reasoning, e.g., deduction, abduction and subsumption, to extract a minimal set of covering models of interpretation from the activity model library based on a set of observed actions, which could explain the observations.

Even though each task can be undertaken in different ways the role of each task is specific and unique. Normally, the first step is to carry out knowledge acquisition, which involves eliciting knowledge from various knowledge sources such as domain experts and activity manuals. The second step is to use various knowledge modelling

techniques and tools to build reusable activity structures. This will be followed by a domain formalization process in which all entities, events and temporal and spatial states pertaining to activities, along with axioms and rules, are formally specified and represented using representation formalism. This process usually generates the domain theory. The following step will be the development of a reasoning engine in terms of knowledge representation formalisms to support the inference. In addition, a number of supportive system components will be developed, which are responsible for aggregating and transforming sensor data into logical terms and formula. With all functional components in place, activity recognition proceeds by passing the logical representation of sensor data onto the reasoning engine. The engine performs logical reasoning, e.g., deduction, abduction or induction, against the domain theory. The reasoning will extract a minimal set of covering models of interpretation from the activity models based on a set of observed actions, which could semantically explain the observations.

There exist a number of logical modelling methods and reasoning algorithms in terms of logical theories and representation formalisms. One thread of work is to map activity recognition to the plan recognition problem in the well-studied artificial intelligence field [66]. The problem of plan recognition can be stated in simple terms as: given a sequence of actions performed by an actor, how to infer the goal pursued by the actor and also to organize the action sequence in terms of a plan structure. Kautz et al. [67] adopted first-order axioms to build a library of hierarchical plans. They proposed a set of hypotheses such as exhaustiveness, disjointedness and minimum cardinality to extract a minimal covering model of interpretation from the hierarchy, based on a set of observed actions. Wobke [68] extends Kautz's work using situation theory to address the different probabilities of inferred plans by defining a partial order relation between plans in terms of levels of plausibility. Bouchard et al. [69] borrow the idea of plan recognition and apply it to activity recognition. They use action Description Logic (DL) to formalize actions and entities and variable states in a smart home to create a domain theory. They model a plan as a sequence of actions and represent it as a lattice structure, which, together with the domain theory, provides an interpretation model for activity recognition. As such, given a sequence of action observations, activity recognition amounts to reasoning against the interpretation model to classify the actions through a lattice structure. It was claimed that the proposed DL models can organize the result of the recognition process into a structured interpretation model in the form of a lattice, rather than a simple disjunction of possible plans without any classification. This minimizes the uncertainty related to the observed actor's activity by bounding the plausible plans set.

Another thread of work is to adopt the highly developed logical theory of actions, such as the Event Calculus (EC) [70], for activity recognition and assistance. The EC formalizes a domain using fluents, events and predicates. Fluents are any properties of the domain that can change over time. Events are the fundamental instrument of change. All changes to a domain are the result of named events. Predicates define relations between events and fluents that specify what happens when and which fluents hold at what times. Predicates also describe the initial situation and the effects of

events. Chen et al. [71] proposed an EC-based framework in which sensor activations are modelled as events, and object states as properties. In addition, they developed a set of high-level logical constructors to model compound activities, i.e. the activities consisting of a number of sequential and/or parallel events. In the framework, an activity trace is simply a sequence of events that happen at different time points. Activity recognition is mapped to deductive reasoning tasks, e.g., temporal projection or explanation, and activity assistance or hazard prevention is mapped to abductive reasoning tasks. The major strength of this work is its capability to address temporal reasoning and the use of compound events to handle uncertainty and flexibility of activity modelling.

Logical activity modelling and reasoning is semantically clear and elegant in computational reasoning. It is also relatively easy to incorporate domain knowledge and heuristics for activity models and data fusion. The weakness of logical approaches is their inability or inherent infeasibility to represent fuzziness and uncertainty. Most of them offer no mechanism for deciding whether one particular model is more effective than another, as long as both of them can be consistent enough to explain the actions observed. There is also a lack of learning ability associated with logic-based methods.

2.4.3 Ontology-Based Approach

Using ontologies for activity recognition is a recent endeavour and has gained growing interest. In the vision-based activity recognition community, researchers have realized that symbolic activity definitions based on the manual specification of a set of rules suffer from limitations in their applicability because the definitions are only deployable to the scenarios for which they have been designed. There is a need for a commonly agreed explicit representation of activity definitions or an ontology. Such ontological activity models are independent of algorithmic choices, thus facilitating portability, interoperability and reuse and sharing of both underlying technologies and systems. Chen et al. [72] propose activity ontologies for analysing social interaction in nursing homes, Hakeem et al. [73] for the classification of meeting videos, and Georis et al. [74] for activities in a bank monitoring setting. To consolidate these efforts and to build a common knowledge base of domain ontologies, a collaborative effort has been made to define ontologies for six major domains of video surveillance. This has led to a video event ontology [75] and the corresponding representation language [76]. For instance, Akdemir [77] used the video event ontologies for activity recognition in both bank and car park monitoring scenarios. In principle, these studies use ontologies to provide common terms as building primitives for activity definitions. Activity recognition is performed using individually preferred algorithms, such as rule-based systems [73] and finite-state machines [77].

In the dense sensing-based activity recognition community, ontologies have been utilised to construct reliable activity models. Such models are able to match different object names with a term in an ontology which is related to a particular activity.

For example, a Mug sensor event could be substituted by a *Cup* event in the activity model "*MakeTea*" as *Mug* and *Cup* can both be used for the "*MakeTea*" activity. This is particularly useful to address model incompleteness and multiple representations of terms. Tapia et al. [64] generate a large object ontology based on the functional similarity between objects from WordNet, which can complete mined activity models from the Web with similar objects. Yamada et al. [78] use ontologies to represent objects in an activity space. By exploiting semantic relationships between things, the reported approach can automatically detect possible activities even given a variety of object characteristics including multiple representation and variability. Similar to vision-based activity recognition, these studies mainly use ontologies to provide activity descriptors for activity definitions. Activity recognition can then be performed based on probabilistic and/or statistical reasoning [64, 78].

Ontology-based modelling and representation have been applied to general ambient assisted living. Latfi et al. [79] propose an ontological architecture of a telehealth-based smart home aiming at high-level intelligent applications for elderly persons suffering from loss of cognitive autonomy. Michael et al. [80] developed an ontology-centred design approach to create a reliable and scalable ambient middleware. Chen et al. [81] pioneered the notion of semantic smart homes in an attempt to leverage the full potential of semantic technologies in the entire lifecycle of assistive living i.e. from data modelling, content generation, activity representation, processing techniques and technologies to assist with the provision and deployment. While these endeavours, together with existing work in both vision- and dense sensing-based activity recognition, provide solid technical underpinnings for ontological data, object, sensor modelling and representation, there is a gap between semantic descriptions of events/objects related to activities and semantic reasoning for activity recognition.

Most works use ontologies either as mapping mechanisms for multiple terms of an object [64] or the categorisation of terms [78] or a common conceptual template for data integration, interoperability and reuse [79]. Activity ontologies which provide an explicit conceptualisation of activities and their interrelationships have only recently emerged and have been used for activity recognition. Chen et al. [82] proposed and developed an ontology-based approach to activity recognition. They constructed context and activity ontologies for explicit domain modelling. Sensor activations over a period of time are mapped to individual contextual information and then fused to build a context at any specific time point. They made use of subsumption reasoning to classify the constructed context based on the activity ontologies, thus inferring the ongoing activity. Ye et al. [83] developed an upper activity ontologies that facilitates to the capturing of domain knowledge to link the meaning implicit in elementary information to higher-level information that is of interest to applications. Riboni et al. [84] investigated the use of activity ontologies, in particular, the new feature of rule representation and rule-based reasoning from OWL2, to model, represent and reason complex activities.

2.5   Discussions on Activity Recognition Approaches

This section presents the comparison of different AR approaches and further discusses the relations between activity recognition and other closely related areas. As activity recognition involves a number of research areas, and each area is itself a research topic with considerable literature. The full reviews of these related areas are beyond the scope of this chapter.

2.5.1   Activity Recognition Approach Comparison

Compared with data-driven and mining-based approaches, ontology-based approaches offer several compelling features: Firstly, ontological ADL models can capture and encode rich domain knowledge and heuristics in a machine-understandable and processable way. This enables knowledge based intelligent processing at a higher degree of automation. Secondly, DL-based descriptive reasoning along a timeline can support incremental progressive activity recognition and assistance as an ADL unfolds. The two levels of abstraction in activity modelling, concepts and instances, also allow coarse-grained and fine-grained activity assistance. Thirdly, as the ADL profile of an inhabitant is essentially a set of instances of ADL concepts, it provides an easy and flexible way to capture a user's activity preferences and styles, thus facilitating personalised ADL assistance. Finally, the unified modelling, representation and reasoning for ADL modelling, recognition and assistance makes it natural and straightforward to support the integration and interoperability between contextual information and ADL recognition. This will support systematic coordinated system development by making use of seamless integration and synergy of a wide range of data and technologies.

Compared with logic-based approaches, ontology-based approaches have the same mechanisms for activity modelling and recognition. However, ontology-based approaches are supported by a solid technological infrastructure that has been developed in the semantic web and ontology-based knowledge engineering communities. Technologies, tools and APIs are available to help carry out each task in the ontology-based approach, e.g., ontology editors for context and activity modelling, web ontology languages for activity representation, semantic repository technologies for large-scale semantic data management and various reasoners for activity inference. This gives ontology-based approaches huge advantage in large-scale adoption, application development and system prototyping.

Logic-based approaches are totally different from data-driven approaches in the way activities are modelled and the mechanisms activities are recognised. They do not require pre-existing large-scale dataset, and activity modelling and recognition is semantically clear and elegant in computational reasoning. It is easy to incorporate domain knowledge and heuristics for activity models and data fusion. The weakness of logical approaches is their inability or inherent infeasibility to represent fuzziness

and uncertainty even though there are recent works trying to integrate fuzzy logics into the logical approaches. Another drawback is that logical activity models are viewed as one-size-fits-all, inflexible for adaption to different users' activity habits. The logical approach, uses logical formalisms, for example event calculus [71] and lattice theory [85], for representing ADL models and conducts activity explanation and predication through deduction or abduction reasoning. Comparing to the above two data-centric approaches, logical approaches are semantically clear in modelling and representation and elegant in inference and reasoning.

A complete comparison between different approaches in terms of a number of criteria is summarised in Tables 2.1 and 2.2. We have collected the experimental results of these surveyed approaches aiming to establish their performance profiles. Initial findings, which are in line with the findings from [86], have found out that the accuracy of different recognition approaches varies dramatically between datasets. The accuracy also varies between individual activities and is affected by the amount of available data, the quality of the labels that were provided for the data, the number of residents in the space that are interacting and performing activities in parallel, and the consistency of the activities themselves. It becomes apparent that the quantitative comparisons of different approaches will only make sense if the experiments are based on the same activities and sensor datasets. Otherwise, the findings may not be applicable to general cases, and even be misleading.

Cook [86] created a single benchmark dataset that contains eleven separate sensor event datasets collected from seven physical testbeds. Using this dataset, a systematic study has been conducted to compare the performance of three activity recognition models: a naïve Bayes classifier (NBC), a hidden Markov model (HMM), and a conditional random field (CRF) model. The result of recognition accuracy using 3-fold cross validation over the dataset is 74.87, 75.05 and 72.16% for the NBC, HMM and CRF respectively.

Table 2.1 The comparison of data-driven approaches

	Generative	Discriminative
Model type	NB, HMM, LDS, DBNs	NN, SVM, CRF, decision tree
Modelling mechanism	(un)supervised learning from datasets	
Activity recognition method	Probabilistic classification	Similarity or rule-based reasoning
Advantage	Modelling uncertainty, temporal information	Modelling uncertainty, temporal information, Heuristics
Disadvantage	"Cold start" problems, lack of reusability and scalability	

Table 2.2 The comparison of knowledge-driven approaches

	Mining	Logic	Ontology	
Model type	HMM, DBN, SVM, CRF, NN	Logic formula, e.g., plans, lattices, event, trees	HMM, DBN, SVM, CRF, NN	Sensor and activity ontology
Modelling mechanism	Information retrieval and analysis	Formal knowledge modelling	(un)supervised learning from datasets	Ontology engineering
Activity recognition method	Generative or discriminative methods	Logical inference, i.e., deduction, induction	Generative or discriminative methods	Semantic reasoning, e.g., subsumption, consistency
Advantage	No "cold start" problems, using multiple data sources	No "cold start" problems, clear semantics on modelling and inference	Shared terms, interoperability and reusability	No "cold start" problems, multiple models, clear semantics on modelling and inference, interoperability and reusability
Disadvantage	The problems as DDA	Weak in handling uncertainty and scalability	The problems as DDA	Weak in handling uncertainty and time

2.5.2 The Influence of Activity Monitoring on Activity Recognition

The outputs of activity sensing, i.e., sensor data, can affect activity recognition in several aspects. Firstly, in a data driven approach, the sensor type can often drive the selection of an appropriate model. Sensors can yield single or multi-dimensional data (e.g., an accelerometer would be multi-dimensional whereas a temperature sensor would be uni-dimensional), and sensors can either give continuous or discrete measurements. The models need to be modified to fit whatever type of sensor data is being used. At the very least, the variable representing each sensor in a data-driven model must match the sensor type in dimensionality and arity. For example, Oliver et al. [57] use a variety of different sensor types, including audio time-of-arrival, continuous and multi-dimensional computer vision measures, and a set of discrete event from mouse and keyboard, as inputs (observations) of a set of HMMs. Liao et al. [52] use continuous 2-dimensional GPS data as input to a CRF. One solution to adapt activity models to sensor types is to include all available sensors in a discriminative or generative model and allow the model itself to choose the most effective ones for any given situation. This is known as sensor selection or active sensing.

Secondly, the complexity of sensor data will determine to some extent the complexity of activity models. In data-driven approaches, sensor data can be directly fed into the activity models, either generative or discriminative, for model training and/or activity inference. Alternatively, sensor data can be pre-processed, e.g., to reduce the complexity of the data, before they are used in model training and activity inference. There is always a trade-off between the complexity of the sensor data in the model, and the complexity of the model. As a general principle the trade-off is always about reducing the complexity of the model as much as possible without sacrificing representation that is necessary for activity recognition.

For knowledge-driven approaches, sensor data do not directly affect activity models and inference. This is because activity models in knowledge-driven approaches are pre-specified based on domain knowledge rather than driven by sensor data. In addition, in knowledge-driven approaches sensor data are always mapped through pre-processing to the values of properties of the formal activity models. As such, the types and complexity of sensor data will only affect the initial conceptualisation of activity models and the complexity of pre-processing but not the model and inference mechanisms.

2.6 Summary

Activity recognition has become the determinant to the success of the new wave of context-aware personalized applications in a number of emerging computing areas, e.g., pervasive computing and smart environments. Synergistic research in various scientific disciplines, e.g., computer vision, artificial intelligence, sensor networks and wireless communications, has resulted in a diversity of approaches and methods to address this issue. In this chapter we present a survey of the state-of-the-art research on sensor-based activity recognition. We first introduce the rationale, methodology, history and evolution of the approach. Then we reviewed the primary approaches and methods in the fields of activity monitoring, modelling and recognition respectively. In particular we identified key characteristics for each individual field and further derived a classification structure to facilitate systematic analysis of the surveyed work. We have conducted in-depth analysis and comparisons of different methods in each category in terms of their robustness to real-world conditions and real-time performance, e.g., applicability, scalability and reusability. The analysis has led to some valuable insights for activity modelling and recognition.

In addition to the extensive review we have discussed emerging research trends associated with activity recognition. One primary direction is complex activity recognition focusing on the underlying modelling, representation and inference of interleaved, concurrent and parallel activities. The other key direction is to improve

reusability, scalability and applicability of existing approaches. Research in this direction has been undertaken in several strands, including multi-level activity modelling, abnormal activity recognition, infrastructure mediated monitoring, and sensor data reuse and repurposing. Another noticeable trend is research on formal activity representation at a higher level of abstraction, e.g., developing dedicated activity representation languages and representing situations and goals. These emerging efforts provide guidance and indication for the future research of activity recognition.

Many research questions have not been touched due to the limited space. For example, we did not elaborate in-depth low-level specific technical issues such as uncertainty, temporal reasoning and sensor data inconsistency. We believe the emerged structure of classification of activity recognition approaches and the comparison of their pros and cons can inform and help interested readers for further exploration

References

1. Mozer MC (1998) The neural network house: an environment that adapts to its inhabitants. In: Proceedings of AAAI spring symposium on intelligent environments
2. Leonhardt U, Magee J (1998) Multi-sensor location tracking. In: Proceedings of the 4th annual ACM/IEEE international conference on mobile computing and networking. ACM, New York, NY, USA, pp 203–214
3. Golding AR, Lesh N (1999) Indoor navigation using a diverse set of cheap, wearable sensors. In: Third international symposium on wearable computers digest of papers, pp 29–36
4. Schmidt A, Beigl M, Gellersen HW (1999) There is more to context than location. Comput Graph
5. Randell C, Muller H (2000) Context awareness by analysing accelerometer data. In: Fourth international symposium on wearable computers digest of papers, pp 175–176
6. Gellersen HW, Schmidt A, Beigl M (2002) Multi-sensor context-awareness in mobile devices and smart artifacts. Mob Netw Appl
7. Van Laerhoven K, Aidoo Ka, Lowette S (2001) Real-time analysis of data from many sensors with neural networks. In: Proceedings of 5th international symposium on wearable computer
8. Foerster F, Fahrenberg J (2000) Motion pattern and posture: correctly assessed by calibrated accelerometers. Behav Res Methods Instruments Comput
9. Laerhoven K, Van Cakmakci O (2000) What shall we teach our pants? In: Fourth international symposium on wearable computers, digest of papers, pp 77–83
10. Lee SW, Mase K (2002) Activity and location recognition using wearable sensors. IEEE Pervasive Comput 1(3):24–32
11. Bao L, Intille SS (2004) Activity recognition from user-annotated acceleration data. In: Ferscha A, Mattern F (eds) Pervasive computing. Springer, Berlin, pp 1–17
12. Patterson DJ, Liao L, Fox D, Kautz H (2003) Inferring high-level behavior from low-level sensors. Presented at the 12 October 2003
13. Nugent CD, Mulvenna MD, Hong X, Devlin S (2009) Experiences in the development of a Smart Lab. Int J Biomed Eng Technol 2:319–331
14. Chan M, Estève D, Escriba C, Campo E (2008) A review of smart homes—present state and future challenges. Comput Methods Programs Biomed 91:55–81
15. Programme A, AAL programme - active assisted living programme - ageing well. http://www.aal-europe.eu/
16. Kern N, Schiele B, Junker H, Lukowicz P, Troster G (2002) Wearable sensing to annotate meeting recordings. In: Proceedings - international symposium on wearable computers, ISWC

17. Lukowicz P, Ward JA, Junker H, Stäger M, Tröster G, Atrash A, Starner T (2004) Recognizing workshop activity using body worn microphones and accelerometers. Presented at the 2004
18. Aggarwal JK, Ryoo MS (2011) Human activity analysis: a review. ACM Comput Surv. 43(3):16
19. Ashbrook D, Starner T (2003) Using GPS to learn significant locations and predict movement across multiple users. Pers Ubiquitous Comput. 7(5):275–286
20. Liao L, Patterson DJ, Fox D, Kautz H (2007) Learning and inferring transportation routines. Artif Intell
21. Sung M, DeVaul R, Jimenez S, Gips J, Pentland A (2004) Shiver motion and core body temperature classification for wearable soldier health monitoring systems. In: Eighth international symposium on wearable computers, 2004. ISWC 2004
22. Harm H, Amft O, Roggen D, Tröster G (2008) SMASH: a distributed sensing and processing garment for the classification of upper body postures. In: Proceedings of the 3rd international ICST conference on body area networks
23. Pantelopoulos A, Bourbakis NG (2010) A survey on wearable sensor-based systems for health monitoring and prognosis
24. Dakopoulos D, Bourbakis NG (2010) Wearable obstacle avoidance electronic travel aids for blind: a survey
25. Yoo J, Cho N, Yoo H-J (2008) Analysis of body sensor network using human body as the channel. In: Proceedings of the ICST 3rd international conference on body area networks. ICST (Institute for computer sciences, social-informatics and telecommunications engineering), ICST, Brussels, Belgium, pp 13:1–13:4
26. Cooper RA, Ding D, Simpson R, Fitzgerald SG, Spaeth DM, Guo S, Koontz AM, Cooper R, Kim J, Boninger ML (2005) Virtual reality and computer-enhanced training applied to wheeled mobility: an overview of work in pittsburgh. Assist Technol 17(2):159–170
27. Au LK, Wu WH, Batalin MA, Stathopoulos T, Kaiser WJ (2008) Demonstration of active guidance with SmartCane. In: 2008 international conference on information processing in sensor networks (ipsn 2008), pp 537–538
28. Kim J, He J, Lyons K, Starner T (2007) The gesture watch: a wireless contact-free gesture based wrist interface. In: Proceedings - international symposium on wearable computers, ISWC
29. Madan A, Caneel R (2004) Towards socially-intelligent wearable networks
30. Wang Q, Timmermans A, Chen W, Jia J, Ding L, Xiong L, Rong J, Markopoulos P (2018) Stroke patients' acceptance of a smart garment for supporting upper extremity rehabilitation. IEEE J Transl Eng Heal Med 6:1–9
31. Wilson D, Atkeson C (2005) Simultaneous tracking and activity recognition (STAR) using many anonymous, binary sensors. In: Proceedings of the third international conference on pervasive computing, (PERVASIVE2005)
32. Wren CR, Tapia EM (2006) Toward scalable activity recognition for sensor networks. In: Lecture notes in computer science (including subseries Lecture notes in artificial intelligence and lecture notes in bioinformatics)
33. Srivastava MB, Muntz R, Potkonjak M (2001) Smart kindergarten: sensor-based wireless networks for smart developmental problem-solving enviroments. In: Proceedings of the 7th annual international conference on mobile computing and networking - MobiCom '01 (2001)
34. Hollosi D, Schröder J, Goetze S, Appell JE (2010) Voice activity detection driven acoustic event classification for monitoring in smart homes. In: 2010 3rd international symposium on applied sciences in biomedical and communication technologies, ISABEL 2010
35. Aipperspach R, Cohen E, Canny J (2006) Modeling human behavior from simple sensors in the home. In: Lecture notes in computer science (including subseries Lecture notes in artificial intelligence and lecture notes in bioinformatics)
36. Philipose M, Fishkin KP, Perkowitz M, Patterson DJ, Fox D, Kautz H, Hähnel D (2004) Inferring activities from interactions with objects
37. Fishkin KP, Philipose M, Rea A (2005) Hands-on RFID: wireless wearables for detecting use of objects. In: Proceedings - international symposium on wearable computers, ISWC
38. Patterson DJ, Fox D, Kautz H, Philipose M (2005) Fine-grained activity recognition by aggregating abstract object usage. In: Proceedings - international symposium on wearable computers, ISWC

39. Hodges MR, Pollack ME (2007) An 'object-use fingerprint': the use of electronic sensors for human identification. In: Proceedings of international conference on ubiquitous computing (UbiComp '07)
40. Buettner M, Prasad R, Philipose M, Wetherall D (2009) Recognizing daily activities with RFID-based sensors. In: Proceedings of the 11th international conference on ubiquitous computing - Ubicomp '09
41. Gu T, Wu Z, Tao X, Pung HK, Lu J (2009) epSICAR: an emerging patterns based approach to sequential, interleaved and concurrent activity recognition. In: 7th annual IEEE international conference on pervasive computing and communications, PerCom 2009
42. Quinn JA, Williams CKI, McIntosh N (2009) Factorial switching linear dynamical systems applied to physiological condition monitoring. IEEE Trans Pattern Anal Mach Intell
43. Horvitz EJ, Breese JS, Heckerman D, Hovel D, Rommelse K (2013) The Lumiere project: Bayesian user modeling for inferring the goals and needs of software users
44. Kautz H, Fox D, Etzioni O, Borriello G, Arnstein L (2002) An overview of the assisted cognition project. In: Proceedings of AAAI
45. Kan P, Huq R, Hoey J, Goetschalckx R, Mihailidis A (2011) The development of an adaptive upper-limb stroke rehabilitation robotic system. J Neuroeng Rehabil 8:33
46. Stikic M, Schiele B (2009) Activity recognition from sparsely labeled data using multi-instance learning. In: Lecture notes in computer science (including subseries Lecture notes in artificial intelligence and lecture notes in bioinformatics)
47. Maurer U, Rowe A, Smailagic A, Siewiorek D (2006) Location and activity recognition using eWatch: a wearable sensor platform. In: Lecture notes in computer science (including subseries Lecture notes in artificial intelligence and lecture notes in bioinformatics)
48. Brdiczka O, Crowley JL, Reignier P (2009) Learning situation models in a smart home. IEEE Trans Syst Man Cybern Part B Cybern 39(1):56–63
49. Chen DT, Yang J, Wactlar H (2005) A study of detecting social interaction with sensors in a nursing home environment
50. Ravi N, Mysore P, Littman ML, Dandekar N (2005) Activity recognition from accelerometer data
51. Vail DL, Veloso MM, Lafferty JD (2007) Conditional random fields for activity recognition. In: Proceedings of the 6th international joint conference on autonomous agents and multiagent systems - AAMAS '07
52. Liao L, Fox D, Kautz H (2007) Hierarchical conditional random fields for GPS-based activity recognition. In: Thrun S, Brooks R, Durrant-Whyte H (eds) Robotics research. Springer, Berlin, pp 487–506
53. Hu DH, Yang Q (2008) CIGAR: concurrent & interleaving goal & activity recognition. In: AAAI conference on artificial intelligence
54. Mahdaviani M, Choudhury T (2007) Fast and scalable training of semi-supervised crfs with application to activity recognition. Adv Neural Inf
55. Guralnik V, Haigh K (2002) Learning models of human behaviour with sequential patterns. In: AAAI workshop on automation as caregiver
56. Modayil J, Levinson R, Harman C (2008) Integrating sensing and cueing for more effective activity reminders. In: AAAI fall symposium AI eldercare new solutions to old problems
57. Oliver N, Garg A, Horvitz E (2004) Layered representations for learning and inferring office activity from multiple sensory channels. Comput Vis Image Underst
58. Pentney W, Philipose M, Bilmes J (2008) Structure learning on large scale common sense statistical models of human state. In: Proceedings of the 23rd national conference on artificial intelligence, vol 3, pp 1389–1395. AAAI Press
59. Wu J, Osuntogun A, Choudhury T, Philipose M, Rehg JM (2007) A scalable approach to activity recognition based on object use. In: Proceedings of the IEEE international conference on computer vision

60. Omar F, Sinn M, Truszkowski J (2010) Comparative analysis of probabilistic models for activity recognition with an instrumented walker. In: Proceedings of the 26th conference on uncertainty in artificial intelligence
61. Sánchez D, Tentori M, Favela J (2008) Activity recognition for the smart hospital. IEEE Intell Syst
62. Wyatt D, Philipose M, Choudhury T (2005) Unsupervised activity recognition using automatically mined common sense. In: Proceedings of 20th national conference on artificial intelligence
63. Perkowitz M, Philipose M, Fishkin K, Patterson DJ (2004) Mining models of human activities from the web. In: Proceedings of the 13th conference on world wide web - WWW '04
64. Tapia EM, Choudhury T, Philipose M (2006) Building reliable activity models using hierarchical shrinkage and mined ontology. In: Lecture notes in computer science (including subseries Lecture notes in artificial intelligence and lecture notes in bioinformatics)
65. Palmes P, Pung HK, Gu T, Xue W, Chen S (2010) Object relevance weight pattern mining for activity recognition and segmentation. Pervasive Mob Comput 6(1):43–57
66. Albrecht D, Zukerman I, Nicholson A (1998) Bayesian models for keyhole plan recognition in an adventure game. User Model User-adapt Interact
67. Kautz Ha (1991) A formal theory of plan recognition and its implementation. Presented at the 1991
68. Wobcke W (2002) Two logical theories of plan recognition. J Log Comput 12(3):371–412
69. Bouchard B, Giroux S, Bouzouane A (2006) A smart home agent for plan recognition of cognitively-impaired patients. J Comput 1(5):53–62
70. Shanahan M (1997) Solving the frame problem: a mathematical investigation of the common sense law of inertia. MIT Press
71. Chen L, Nugent C, Mulvenna M, Finlay D, Hong X, Poland M (2008) A logical framework for behaviour reasoning and assistance in a smart home. Lecture notes in computer science (including subseries Lecture notes in artificial intelligence and lecture notes in bioinformatics)
72. Chen D, Yang J, Wactlar HD (2004) Towards automatic analysis of social interaction patterns in a nursing home environment from video. In: Proceedings of the 6th ACM SIGMM international workshop on multimedia information retrieval - MIR '04
73. Hakeem A, Shah M (2004) Ontology and taxonomy collaborated framework for meeting classification. In: Proceedings - international conference on pattern recognition
74. Georis B (2004) A video interpretation platform applied to bank agency monitoring. In: Intelligent distributed surveillance systems (IDSS-04) (2004)
75. Nevatia R, Hobbs J, Bolles B (2004) An ontology for video event representation. In: IEEE computer society conference on computer vision and pattern recognition workshops
76. François ARJ, Nevatia R, Hobbs J, Bolles RC (2005) VERL: an ontology framework for representing and annotating video events. IEEE Multimed 12(4):76–86
77. Akdemir U, Turaga P, Chellappa R (2008) An ontology based approach for activity recognition from video. In: Proceeding of the 16th ACM international conference on multimedia - MM '08
78. Yamada N, Sakamoto K, Kunito G, Isoda Y, Yamazaki K, Tanaka S (2007) Applying ontology and probabilistic model to human activity recognition from surrounding things. IPSJ Digit Cour
79. Latfi F, Lefebvre B, Descheneaux C (2007) Ontology-based management of the telehealth smart home, dedicated to elderly in loss of cognitive autonomy. In: CEUR workshop proceedings
80. Klein M, Schmidt A, Lauer R (2007) Ontology-centred design of an ambient middleware for assisted living: the case of SOPRANO. Context
81. Chen L, Nugent C, Mulvenna M, Finlay D, Hong X (2009) Semantic smart homes: towards knowledge rich assisted living environments. Stud Comput Intell
82. Chen L, Nugent CD, Wang H (2012) A knowledge-driven approach to activity recognition in smart homes. IEEE Trans Knowl Data Eng 24(6):961–974
83. Ye J, Stevenson G, Dobson S (2011) A top-level ontology for smart environments. Pervasive Mob Comput 7(3):359–378

84. Riboni D, Bettini C (2011) OWL 2 modeling and reasoning with complex human activities. Pervasive Mob Comput 7(3):379–395
85. Preuveneers D, den Bergh J, Wagelaar D, Georges A, Rigole P, Clerckx T, Berbers Y, Coninx K, Jonckers V, De Bosschere K (2004) Towards an extensible context ontology for ambient intelligence. In: Markopoulos P, Eggen B, Aarts E, Crowley JL (eds) Ambient intelligence. Springer, Berlin, pp 148–159
86. Cook DJ (2012) Learning setting-generalized activity models for smart spaces. IEEE Intell Syst 2010(99):1