

# A Survey on Context-Aware Sensing for Body Sensor Networks

Barbara T. Korel<sup>1</sup>, Simon G. M. Koo<sup>2</sup>

<sup>1</sup>Department of Computer Science, Brown University, Providence, USA

<sup>2</sup>Department of Mathematics and Computer Science, University of San Diego, San Diego, USA

E-mail: [bkorel@cs.brown.edu](mailto:bkorel@cs.brown.edu), [koo@sandiego.edu](mailto:koo@sandiego.edu)

Received May 14, 2010; revised May 25, 2010; accepted May 27, 2010

## Abstract

Context awareness in Body Sensor Networks (BSNs) has the significance of associating physiological user activity and the environment to the sensed signals of the user. The context information derived from a BSN can be used in pervasive healthcare monitoring for relating importance to events and specifically for accurate episode detection. In this paper, we address the issue of context-aware sensing in BSNs, and survey different techniques for deducing context awareness.

**Keywords:** Context Awareness, Body Sensor Networks, Artificial Neural Networks, Bayesian Networks, Hidden Markov Models

## 1. Introduction

Context is defined as “any information that can be used to characterize the situation of an entity, where an entity can be a person, place or physical object” [1]. Context awareness can then be defined as detecting a user’s internal or external state. Context-aware computing describes the situation of a wearable or mobile computer being aware of the user’s state and surroundings, and modifying its behavior based on this information [2]. Context awareness plays a significant role in Body Sensor Networks (BSNs) because it interprets physical and biochemical signals coming from the BSN, based on information regarding the current state of the user and the state of the environment. Context-aware sensing is an integral part of the BSN design to achieve the ultimate goal of long-term pervasive health care monitoring.

There are three main approaches that have been applied to deduce context in a sensor network: Artificial Neural Networks, Bayesian Networks and Hidden Markov Models. Research in context awareness or activity recognition using these methods has primarily been done in wireless sensor networks or wearable sensor networks, so the application of context-aware sensing in BSNs is still new and faces many technical challenges. This paper will address some of these issues raised, describe the characteristics of each method, and discuss how these

This work is supported in part by the Faculty Research Grant, College of Arts and Sciences, University of San Diego.

algorithms handle the challenges that need to be faced in context sensing for BSNs. As a note of credit, this paper is inspired and largely based upon [3].

## 2. Context Awareness in BSNs

Wireless medical body sensor devices, either implantable or wearable, are used to monitor a patient’s physiological state including EKG, heart rate, blood pressure, oxygen saturation and sweat volume/rate. The wireless BSN framework is designed to provide such pervasive monitoring of the human body; this ultimately has a huge impact on medical healthcare and monitoring vital signs of elderly patients or patients with chronic cardiac disease. BSNs present a method to continuously monitor physiological parameters to detect life threatening abnormalities that could lead to mortality. In addition to a patient’s vital signs, a person is physiologically very sensitive to external context or environmental change. Such contextual factors include the person’s activity, current temperature of the outside environment, time of day, etc. For instance, if a body sensor detects a rapid increase in a patient’s heart rate, the patient might not be experiencing a cardiac episode, but rather undergoing a change in his physical activity such as jogging. By incorporating context awareness into the BSN, environmental factors and the state of the patient can be evaluated. Ultimately, changes in the physiological state of the body can be rationalized according to the events that triggered such

changes.

There are various algorithms for context-aware sensing that can deduce context in a BSN; each has different characteristics and accomplishes different tasks. In many applications studied, these approaches are actually used in combination with one another to achieve context from the environment. The first step in achieving context awareness in a sensor network is to gather the low-level sensor readings from all sensor nodes; these data readings will always constitute as the input for context sensing. It is often beneficial for the input data to be ordered in some way, thus the data is typically clustered into subgroups such that distance is small among data entries in the same cluster and distance is large among data entries from different clusters [4]. This is accomplished by a clustering algorithm. To actually achieve context from the sensor nodes, the input or clustered data must be associated with a context. In the domain of sensor networks, this process is known as classification, which associates the input vectors to a context profile through the means of user labels. Some classification algorithms are only able to recognize context at a given instance in time but not continuously. So a supervising layer is introduced on top of the classification layer to extract constant recognition of context. Algorithms performing at the supervising layer are able to classify context transitions and more closely model context events occurring in natural human behavior.

### 3. Challenges of Context-Aware Sensing

In practice, BSNs for pervasive healthcare monitoring will result in network applications operating in a variety of different environments including a hospital operating room, an elderly health clinic or a personal home setting. Each of these environments varies substantially from one another and yet the BSN framework must be adaptable and distributed to accommodate for such different settings. Due to the diversity of context-aware environments, the range of physiological conditions a patient may experience, and the dynamic nature of BSNs themselves, many challenges arise for context-aware sensing. Specific issues include overcoming sensor noise, node failure and motion artifact in the network, integrating multi-sensory data, allowing for smooth context recognition, providing long term/continuous usage of the application, appropriately structuring the network in terms of number of sensors, and selecting relevant features in the BSN. These are the challenges context-aware sensing faces and the various methods, discussed in later sections, handle some of these issues in various ways.

#### 3.1. Noise Resilience and Detection

Noise in a sensor network may result from sensor noise,

node failure or motion artifact [3]. The presence of noise in the network from any of these sources may introduce significant errors into the input data of the sensor network. This data may contain missing sensor information, malicious sensor readings or uncertain information; the resultant input vectors will not contain an accurate representation of the sensor readings. For context sensing in BSNs, this could have unfortunate consequences because in pervasive healthcare monitoring, detrimental actions are usually taken based upon the sensed values. Specifically the cost of any “unclean” data can be very significant since it is used for critical decisions [5]. The quality of input sensor readings is crucial and the presence of noise degrades this quality of data obtained from the sensor network. Thus a context sensing algorithm must be able to detect such malicious noise and reduce the effect of noise in sensor data to appropriately model the network.

#### 3.2. Introduction of Smoothness Constraint

Human activity innately involves body movement that is continuous in nature; if such a smoothness constraint were enforced in BSNs, context could be recognized with a higher accuracy based on natural human behavior [6]. Typically context is classified at a given instance in time, but introducing a smoothness constraint means the system must sense the transitions between individual context or sequences of context. Context sensing in BSNs must be able to capture such transitions to accurately recognize context from the continuous flow of human movement in time.

#### 3.3. Adaptive On-Line Learning

Recognizing context in the real-world domain of BSNs is a function that needs to remain flexible as new context may be continuously added to the system and old context may no longer be perceived. Thus the system needs to remain adaptive to learn new context online from the sensor network. Since real BSN applications will function over longer periods of time, it is important that the system not only be able to learn new context as it is presented to the system, but also to not forget previously encountered context that was learnt, so it doesn't have to re-learn such context again.

#### 3.4. Input Data Dimensionality

A large number of sensors in the BSN may be necessary in order to achieve accurate context recognition of a patient's state/activity or if the system is to recognize a large number of different contexts. By adding more sensors to the network, context recognition can be achieved with a higher accuracy [7]. However, two significant

problems may arise from a high dimensionality of input sensors. First, a substantial burden may be placed on the power consumption and bandwidth of the system as more sensors are added to the network. This is rather an issue of reducing the transmission range and required bandwidth; clustering data transmission will also reduce power consumption [8].

The second problem is known as the curse of dimensionality: as the number of sensor inputs increases, the learning rate of the algorithm significantly slows down. For a BSN to maintain good performance, a system composed of a large number of sensors should not slow down or decrease fault tolerance.

### 3.5. Feature Selection

If only the relevant sensor readings were applied to the current context, irrelevant or redundant sensors could be filtered out. This is beneficial because typically the number of features in the BSN is numerous, especially if there is a large number of sensors. However, only a small subset of those features is necessary or relevant to recognize the context [9]. **Figure 1** illustrates how context is extracted in a sensor network using feature selection.

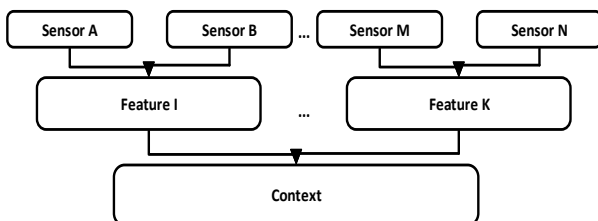
Thus if the input data that is not useful in the decision process of context recognition were not sent across the wireless network, the dimensionality of data will be reduced. Essentially this allows for a decrease in data transmission (implying less power consumption and bandwidth) and efficient data mining of the BSN as only relevant information is used in the sensor network [10].

## 4. Techniques for Context Recognition

This section presents a survey of various techniques that have been used in context-aware sensing. The work by Yang [3] served as a major inspiration to our study.

### 4.1. Artificial Neural Networks

The Artificial Neural Network (ANN) is used as a solid clustering algorithm for context awareness in sensor networks. It is based on the biological nervous system of the



**Figure 1. Context extraction in a sensor network using feature selection.**

brain that consists of a large number of small and simple interconnected components: neurons. Each neuron can perform its own set of computations, yet the network is capable of performing powerful computations by combining the limited processing power of each element [7].

For sensor networks used in practice, the low-level sensors will produce some level of noise no matter what. One of the key advantages of using ANNs is that they are still able to perform well despite the presence of this inevitable noise coming from data sensors. Another beneficial characteristic of neural networks includes unsupervised training of the input data, that is, the user of a wearable computer does not need to spend much time training it and the context learning for the system is not limited to just the training phase. So as the user transitions from context to context, the algorithm should learn autonomously the context from the new input it receives by recalculating its internal representation of the context (known as on-line adaptation) [11]. Thus because the data is able to approximate itself, the neural network can feasibly add new context to the system when necessary, without intervention from the user.

The following discusses two types of ANNs: Kohonen Self-Organizing Map (KSOM) and KSOM with k-means clustering. **Table 1** presents the pros and cons of using ANNs for context awareness and **Table 2** summarizes some studies and applications of ANNs in sensor networks.

#### 4.1.1. Kohonen Self-Organizing Maps

The Kohonen Self-Organizing Map (KSOM) is a type of unsupervised neural network which is used to cluster the input vectors (low-level sensor readings from nodes) to a discrete output space that is in the form of a grid-like map. Just as described by ANNs, the outcome is similar signals are mapped close to each other on the map and dissimilar signals are mapped at greater distances from each other [18].

**Algorithm:** The structure of the KSOM consists of an input layer and an output layer: the input layer is essentially the input vector of data and each input node is assigned a map-unit to introduce order among the input vectors. The output layer is a grid of interconnected neurons, usually as a one or two-dimensional array. Each neuron in the output layer is connected to every single neuron from the input layer and this connection is assigned a particular weight. In addition, every neuron is also connected with its nearest neighbor nodes on the grid-map. **Figure 2** depicts the structure of a bi-dimensional KSOM.

The KSOM is a competitive network in that each unit in the output layer competes with the other output units for a particular kind of input. So when a new input value is presented to the KSOM, the input vector is compared to each output neuron's weight vector. The output neuron that has a weight vector closest to the input vector is se-

**Table 1. Benefits and drawbacks of the ANN approach to context awareness.**

Technique	Advantages	Limitations
ANN KSOM	1) Provides a means of visualizing the behavior of the network. 2) Unsupervised training of input data, <i>i.e.</i> , doesn't require user participation. 3) Provides an efficient means to cluster data.	1) <i>Stability-plasticity dilemma</i> : KSOM does not remain adaptive over time. 2) <i>Curse of dimensionality</i> : high dimensionality of input data results in a slow and less fault tolerant algorithm.
ANN KSOM with k-means clustering	Overcomes stability-plasticity dilemma so it remains adaptive and stable over time.	1) Requires labeled input vectors, meaning 2) User participation is necessary.

**Table 2. Techniques in context-aware sensing.**

	Application	Experiment	Results
Schmidt <i>et al.</i> (1999) [12]	Presents a layered real-time architecture for context-aware adaptation based on redundant collections of low-level sensors; the context is derived using KSOM.	A prototype board consisting of 8 sensors, a PDA and a mobile phone were used to demonstrate situational awareness.	Experiments show that it is feasible to recognize context using sensors and that context information can be used to create new interaction metaphors.
Lagerholm <i>et al.</i> (2000) [13]	Uses unsupervised self-organizing NNs to cluster ECG signals into 25 groups.	Cluster ECG complexes into classes which are not predefined.	Using the MIT-BIH arrhythmia database, the resulting KSOM clusters exhibit a very low degree of misclassification (1.5%).
van Laerhoven <i>et al.</i> (2000) [14]	Shows an integrated approach using KSOM for clustering, along with K-nearest neighbor for classifying different activities.	A pair of pants with accelerometers, connected to a laptop to interpret raw sensor data, are used to recognize activities such as walking, sitting, climbing stairs, etc.	The user has the ability to decide what activities are learned at what time, while the system remains autonomous enough such that the interaction is kept very minimal.
van Laerhoven <i>et al.</i> (2001) [7]	Shows that Neural Network is an ideal algorithm to analyze data coming from a large number of small and simple sensors; specifically KSOM with k-means clustering.	A wearable system consisting of several simple sensors used to learn different, simple activities like sitting, standing and walking; also used to automatically start processes or tasks depending on the current context.	Determining different context a wearable computer can encounter, by merely labeling them as they occur, is still difficult to realize without setting harsh constraints on system, usually in terms of the available context.
van Laerhoven <i>et al.</i> (2001) [11]	Uses KSOM with k-means clustering for classification of incoming sensor data in a real-time fashion.	Two accelerometers placed above the knee to recognize simple, everyday activities such as sitting, standing, walking, running and bicycling.	KSOM can be very unstable in initial phases; k-means algorithm added to KSOM to overcome overwriting previous inputs and to create a stable topological mapping of sensor data.
Simelius <i>et al.</i> (2003) [15]	Presents SOM for spatiotemporal analysis and classification of body surface potential mapping (BSPM) data.	86 cardiac depolarization (QRS) sequences paced by a catheter in 18 patients, in which spatial BSPM distributions at every 5ms over the QRS complex were presented to an untrained SOM.	This method has potential for determining abnormal ventricular activity.
Gao <i>et al.</i> (2004) [16]	Presents a diagnostic system for cardiac arrhythmias from ECG data, using an ANN classifier (based on a Bayesian framework).	Bayesian ANN-based arrhythmia diagnostic system determines a patient's current condition in real-time, using ECG signals.	At least 75% prediction accuracy for classification in both the training and test phases. Greater than 90% false rate (cry wolf dilemma of medical monitoring devices—a false alarm is raised yet patient needs no attention) prediction accuracy in both phases.
Thiemjarus <i>et al.</i> (2006) [17]	Proposes a spatio-temporal SOM that minimizes the number of neurons involved while maintaining a high accuracy in class separation for both static and dynamic activities.	Four accelerometers were placed on left and right ankles and legs for a simple physical exercise sensing experiment involving sitting, standing, steps, dem-plie, galloping, skipping, etc.	Using standard SOMs, the average performance was about 58%, and an increase in the number of neurons from 100 to 400 did not make a noticeable difference. The use of spatio-temporal SOM with a relatively small number of neurons shows a great improvement in performance.

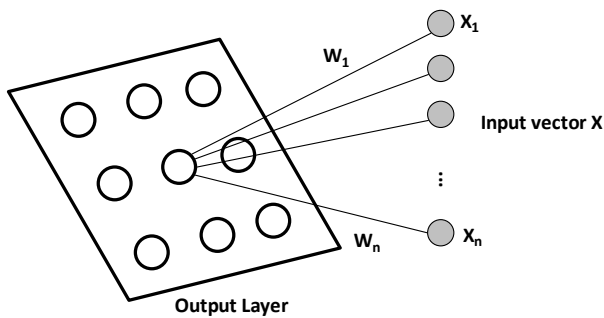


Figure 2. The structure of a bi-dimensional KSOM.

lected as the winner node and is able to adapt itself more towards the input. This means the winning output node updates its weight vector to more closely reflect the values of the input vector. To introduce topological ordering among related units, the neighboring nodes of the winning node are also permitted to update their weight vectors towards the input vector, but to a lesser degree. After the first iteration of this algorithm with the given input data, errors will typically exist when new signal readings are introduced to the mapping [18]. However, after only a few iterations of this algorithm, the data organizes itself in a structured and topological way such that similar sensor signals activate neighboring units and different signals activate different neurons [14]. Thus the KSOM clusters  $n$ -dimensional input data from the sensors into an array of neurons in an adaptive (meaning the neurons in the grid “learn” to respond better for particular input), and unsupervised fashion.

**Application to Context Awareness:** After the input from sensor readings has been clustered by a clustering algorithm, it becomes significantly easier to process or classify the data. In terms of context awareness, KSOM is in general a universal approach to process sensor data because it does not require a priori knowledge of the context and is able to perform learning without explicit user supervision [14]. Thus using the KSOM to topologically map sensor data is appropriate in applications which may not contain labeled training data or where activities are not well-defined. In terms of BSNs, this is beneficial because it enables the system to not only detect context that has not yet been defined by the user, but it also allows the system to capture context that is unpredictable and randomly appears in the system.

Any system in the real-world that is to deduce context awareness from body sensors must have the requirement of remaining adaptive over time. KSOMs have the capability to be adaptive; however this comes with its limitations. The KSOM algorithm starts out highly adaptive with a large learning rate, and then over time becomes fixed so that it is no longer capable of learning anymore [14]. This problem is known as the *stability-plasticity*

*dilemma* and is one of the main drawbacks to KSOMs such that they are not adaptive over time. If the system were designed to remain flexible, then overwriting previously learned instances would occur as neurons belonging to a learned cluster would gradually change to other clusters. This is true because the KSOM has a fixed structure and cannot infinitely grow over time, (*i.e.*, it has a limited number of context it can cluster, so the number of clusters remains the same while the neurons adapt to different clusters within the same set of clusters). Overall the stability-plasticity dilemma prevents a system from using KSOM for long-term use since the system either would gradually become fixed or would become unstable since previously learned context would be forgotten [7].

**Limitations:** A BSN should not be limited by the number of sensors or the number of inputs in the network. In practice, if the number of input increases in the system, it should be able to handle a large number of input data and not slow down. Unfortunately, KSOM suffers from this curse of dimensionality and causes a high dimensionality of input data to be a problem for this algorithm [14]. As the number of sensors increases in the system, the clustering algorithm must map the high input space to a large output space and uses a lot of resources to do so, resulting in a slow and less fault tolerant algorithm. This problem becomes especially bad if there is a lot of noise or irrelevant sensor nodes in the system, in which the algorithm maps irrelevant context to its output space and wastes many resources in the process. Thus the KSOM proves to have limitations in the real-world domain where the input space could be very large and irrelevant context is undoubtedly present, causing a serious digress in maintaining performance speed and fault tolerance.

#### 4.1.2. KSOM with K-Means Clustering

The traditional KSOM has the disadvantage of “unlearning” context or overwriting prototypes on the map if the algorithm is to remain adaptive, known as the *stability-plasticity dilemma*. This is a major shortcoming for context awareness in BSNs since the system should be able to learn new context over a long period of time and not forget old context. This problem can be overcome by introducing a k-means clustering algorithm to the KSOM. The KSOM still clusters the sensor input and preserves map topology; the k-means clustering algorithm then clusters labeled input vectors a second time and adds a second layer to the structure [11]. **Figure 3** illustrates KSOM with the k-means clustering layer.

**Advantages over KSOM:** The k-means clustering algorithm has two main differences from the traditional KSOM: it is not topology preserving and it requires labeled input vectors (meaning the user needs to participate in the training phase to label incoming sensor data with a context description) [11]. With just the KSOM, clusters are unlabeled making the algorithm unaware of

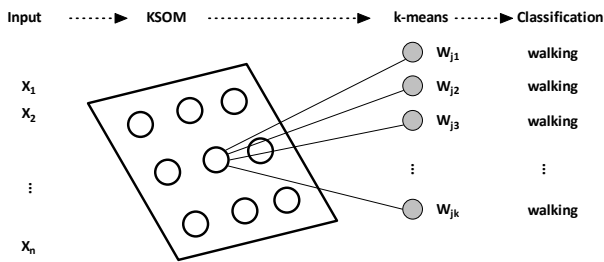


Figure 3. KSOM with k-means clustering layers.

relevant context associated with clusters. However, in k-means the user specifies the labels making the algorithm aware of existing context so it knows not to overwrite clusters which contain relevant context. Also, because k-means is not topology preserving, the hierarchical structure allows the k-means sub-clusters to preserve already clustered data when KSOMs' topological mapping begins to overwrite previously learned prototypes. Overall, the addition of a second layer with k-means clustering allows a context-aware system in BSNs to function over a long period of time while remaining adaptive and stable.

## 4.2. Bayesian Networks

Bayesian Networks (BNs) are an appropriate method for deducing context awareness by classifying context from the associated sensor readings in the system. The Bayesian Network is a form of a graphical probabilistic model in which the structure of a BN is a directed acyclic graph. The nodes in the graph signify random variables and the directed arcs between nodes represent their causal dependencies. Thus the set of random variables is the domain of interest and all of the direct causal or influential relationships are encoded by the arcs. BNs follow an independence assumption that every node in the graph is strictly independent of any other variable except its descendants. Theoretically, simple BNs are considered to be ideal, such that they obtain the highest accuracy only when this independence assumption holds. This independence assumption along with the graphic structure representing unambiguous interdependent relationships allows for an important feature of BNs: representation of joint probability distributions.

Bayesian networks can also be dynamic, which means that the BN determines the activity being performed based on variables from the sensory data [19]. Essentially, dynamic BNs represent sequences of variables (*i.e.* nodes in the graph of a BN). In a dynamic BN, these sequences are activities that are being performed, and these activities are determined by certain variables. The activities are viewed as hidden variables, and the observed variables include the set of objects seen and the time elapsed. Using the observed variables (*i.e.*, time

elapsed), it becomes possible to probabilistically estimate the most likely activities and their intensity from sensor data by using Bayes filtering. A sequential Monte Carlo approximation is also used to solve for the most likely activities based on sensor data.

Along with recognizing what activity the person is performing, it is also important to figure out the person's emotional state [20]. However, deciphering a person's emotional state through sensory data is extremely difficult, as it is not always the case that the human perfectly knows someone else's emotion. Nevertheless, researching general emotions that accompany certain activities helps to figure out someone's current state. It should be noted that this may not always be accurate. Using sensory data, the BSN can detect emotions for every instance. By recognizing human emotion in these instances, the BSN can help to assess stress, anger, and other emotions that pertain to the individual's health. Furthermore, relating stress to the use of a product may give significant information to help developers redesign and improve their products to better fit humans needs.

Another technique that can be used to determine human activities is a decision tree [21]. With this method, there are several possible activities that can occur in one given place. The activity is then determined by sensing what the human is doing in the current environment. Essentially, there is a preset number of possible activities given in a certain situation, but sensory data determines which one of these activities is actually taking place.

The mixture model method brings another approach to determining human activities [22]. Mixture models cluster observations into event types, and activities are considered human behaviors. After mixture models organize observations into clusters, a density function, which can be found in [9], is applied to check the significance of the clusters. Density represents a calculation of the consistency of a given activity. The higher the density, the more significant a cluster is. Therefore, activities with the highest density calculations are the most commonly occurring activities, and those with low density calculations are classified as random events. Based on this information, it is easier for the BSN to determine what event is most likely occurring.

The following sections describe naïve Bayes Classifiers, and BNs with hidden nodes. **Table 3** discusses the advantages and limitations of BNs for context-aware sensing, and **Table 4** summarizes various applications of BNs in context-aware sensing.

### 4.2.1. Naïve Bayes Classifiers

A naïve Bayes classifier is a probabilistic classifier adhering to Bayes' rule and in context awareness is used for classification. More specifically, activity recognition may be reduced to a classification problem where classes correspond to activities and Bayes classifiers predict the activity labels after training examples are generated.

**Table 3. Benefits and drawbacks of the Bayesian Network approach to context awareness.**

Technique	Advantages	Limitations
Naïve Bayes Classifiers	<ol style="list-style-type: none"> <li>1) Helpful in predicting consequences of interaction within the domain, since BNs provide the ability to gain an understanding of the problem domain at hand by the existent relationships.</li> <li>2) Computationally efficient, which means viable for real-time recognition.</li> <li>3) Noise resilient. It does not model ambiguous or noisy information from multiple sensors; handles classifying context where missing or incomplete data exists.</li> </ol>	<ol style="list-style-type: none"> <li>1) If there is confusion in labeled training data between similar activities, or if there is an extremely low number of training data, activity recognition suffers greatly.</li> <li>2) Requires labeled training data, thus needs user participation.</li> <li>3) Large scope of real-world domains may cause naïve Bayes to perform poorly since not all objects and relations may be represented.</li> <li>4) Problem of genericity in recognition due to ambiguity of some features referring to multiple context.</li> <li>5) Enforces mutual exclusivity so when two different activities occur simultaneously, detection of one could preclude detection of the other.</li> </ol>
Feature Selection and Smoothness Constraint	<ol style="list-style-type: none"> <li>1) Additional noise resilience and detection, which in turn increases model accuracy.</li> <li>2) Allows for feature selection.</li> <li>3) Neutralizes redundancy in the network.</li> <li>4) Smoothness constraint introduced to model continuous human movement.</li> <li>5) Decreases power consumption of network due to feature selection.</li> <li>6) Allows for distributed computing across network to local subnets which reduces stress on central processor in terms of bandwidth and computational load.</li> </ol>	<ol style="list-style-type: none"> <li>1) The distinction between very similar context may be blurred or lost.</li> <li>2) Requires user participation.</li> </ol>

**Table 4. Examples of some BN techniques in context-aware sensing.**

	Application	Experiment	Results
Madabhushi <i>et al.</i> (1999) [23]	Incorporates a Bayesian framework to human activity recognition in order to automatically identify human action.	Classifies ten different human actions from visual information by tracking the position of the head in pictures.	The system had a success rate of 80% for recognizing activity. The system has limitations such as only recognizing one action in a sequence and not performing in real time.
Korpipaa <i>et al.</i> (2003) [24]	Applies naïve Bayesian networks to classify the context of a mobile device user in her normal daily activities.	The classification was based mainly on audio features measured in a home scenario.	Situations can be extracted fairly well, but most of the context is likely to be valid only in a restricted scenario; naïve Bayes framework is feasible for context recognition.
Tapia <i>et al.</i> (2004) [25]	Proposes a system based on a naïve Bayesian framework for recognizing activities in the home setting.	Uses a set of small and simple state-change sensors that can be quickly and ubiquitously installed in the home environment to recognize activities.	Results from a small dataset show that it is possible to recognize activities such as toileting, bathing and grooming with detection accuracies ranging from 25% to 89%.
Elnahrawy <i>et al.</i> (2004) [26]	Proposes a technique based on Bayesian classifiers for modeling and learning statistical contextual information in sensor networks.	Analyzes the approach in two applications, tracking and monitoring. Introduces applications of the model in outlier detection, approximation of missing values and sampling.	Once the contextual information is learned, these applications reduce to an inference problem. Evaluations show the applicability and a good performance of the approach.

Thus naïve Bayes classifiers require labeled training data to recognize clearly defined activities, which has the downside of requiring more effort in the data recording phase [18]. However, with the given training sample, naïve Bayes classifiers are able to optimally predict a

class of examples that have not been previous seen by the system [27]. Generally speaking the theory of probability provides a solid ground for the task of classification [24]; since the naïve Bayes classifier is a probabilistic induction algorithm, this is an approach to classifica-

tion that performs with high accuracy and an attainable recognition rate for activity recognition in specific domains.

As mentioned previously the Bayesian classifier is thought to perform optimally when it adheres to the independence assumption in which there are no dependencies between attributes. However, in many domains including the real-world, this assumption does not and cannot hold true; thus it seems that either this assumption is undermined or the accuracy of the Bayesian classifier is not in fact optimal. However in [27], Domingos and Pazzani showed that the Bayesian classifier is optimal despite strong attribute dependencies existing in the system. This is partially due to Bayes classifiers not depending on the original independence assumption to perform at their best. Another explanation could be that the high bias that exists due to the strong independence assumption could be neutralized by the low variance of the classifier [25].

**Limitations:** Activity recognition accuracy suffers based on naïve Bayes when either there is confusion in the labeled training examples between very similar activities or there is an extremely low number of training examples for an activity. In many applications which have implemented naïve Bayes for context awareness in wireless sensor networks, testing was performed in a restricted scenario; had testing been done in a real-world scenario the accuracy rates would most likely have been lower. This indicates that naïve Bayes classifiers may perform poorly in real domains. One reason could be that in the real-world, naïve Bayes classifiers hold the assumption that all attributes that influence a classification decision are observable and represented [25], which may be the case for a limited test scenario but not in the scope of the real-world. In the real-world domain there are many different kinds of objects, classes of objects and numerous relations among them. BNs are still limited in not being able to exhaustively represent all the objects and relations that exist in the real-world [28]. Problems of genericity in recognition will also arise for naïve Bayes classifiers due to the ambiguity of some features referring to multiple contexts [24]. An additional disadvantage of the naïve classifier is it enforces mutual exclusivity, thus when two different activities occur simultaneously, detection of one could preclude the detection of the other [25].

**Advantages:** Key advantages of naïve Bayes classifiers include providing a noise resilient classification framework by not modeling ambiguous or noisy information from multiple sensors in the network. In addition they appropriately handle classifying situations where missing or incomplete data exists. BNs have also been shown to be computationally efficient, making them viable for real-time recognition of context (an important criterion for body sensor networks) [24]. Lastly naïve Bayesian networks provide the ability to gain an overall

understanding of the problem domain at hand from learning the causal relationships that exist. This can then be used for predicting the consequences of interaction within the domain [29].

#### 4.2.2. Hidden Nodes

In a BN, there is a considerable amount of dependency between parent and child nodes which violates the independence assumption that BNs are based on. This dependency may increase when redundant nodes are added to the network to overcome motion artifact and sensor failure [6]. Hidden nodes may be added to the BN as unobserved variables to represent the dependencies among children. They effectively compensate the extra weight dependant children add to the network, neutralizing redundancy in the network. **Figure 4** shows the structure of a BN with hidden nodes.

**Advantages of Hidden Nodes:** By inserting hidden nodes to the BNs, subnets are formed (representing the redundant nodes) to increase robustness and benefit the network in many ways. First the subnets provide additional noise resilience by filtering out noise present within each subnet, in turn increasing model accuracy [8]. Additionally, hidden nodes are able to detect node failures by identifying an asynchronous child-parent dependency, thus providing noise detection [6]. The capability of BNs with hidden nodes to detect noise and accurately classify context despite sensor noise is crucial in BSNs where detrimental actions are taken based on the context recognized. The smoothness constraint necessary in modeling continuous human movement is possible using BNs with hidden nodes. This can be achieved by adding a fixed size temporal window over instantaneous model beliefs of the network [6]. Inserting hidden nodes to indicate redundant sensors is a form of feature selection in which irrelevant sensor data is not sent across the network, which has the advantage of decreasing net power consumption. Finally, stress on the central processor of the network can be diminished in terms of bandwidth and computational load by distributing computations across the network to the local subnets [8].

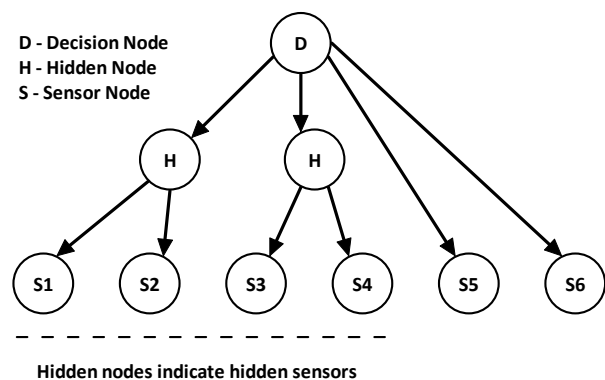


Figure 4. Bayesian network with hidden nodes.



### 4.3. Hidden Markov Models

For context sensing in a body sensor system to be applicable in the real-world, context needs to be continually recognized throughout a duration of time and not just at exact instances in time (as the case of BNs). Hidden Markov Models (HMMs) are introduced at the supervising layer to achieve a model of context transition. The use of HMMs is an approach to context recognition that can more accurately model human behavior since the system is capable of recognizing sequences of activities.

A feature selection technique that can be used for context-aware sensing is the HMM. HMMs transform a set of sequences of different lengths to a feature space, where clustering can then be performed [30]. In this feature space, each sequence is represented as a vector of distances. Using HMMs, each sensor output is represented by its distance to HMMs trained for the whole set of sensor outputs instead of comparing sensor outputs directly as time series of different lengths. This distance refers to how likely a certain event is to be predicted by that HMM and is described by the log likelihood. If a sensor's output is highly correlated to the trained HMM, the values of the associated log likelihood would also be high.

Layered HMMs (LHMM) are separate layers of HMMs connected according to their inferential results [31]. In this way, each level of the LHMM hierarchy can be trained in a different way. Instead of having a whole HMM responsible for all actions, LHMM allow for different layers to be responsible for detecting certain actions. There are two ways to do inference with LHMMs. One way is to select the model in which the highest likelihood is selected; this is called *maxbelief*. This information is made available as input to the HMM at the next level of the hierarchy. In the distributional approach, the full probability distribution over the models is passed to the higher-level HMMs.

**Algorithm:** HMMs are probabilistic models used to represent non-deterministic processes and consist of states, actions and observations. The outcome or observation of a state is determined by the conditional probability distribution of the state and is based on the Markov property, where the current state of the environment depends solely on the intermediate previous state and the associated action. In HMMs the state sequence is hidden and only the observations are visible. The transition from one state to another is an action, labeled with the probability that the transition from one context to another will occur. In context awareness sensing, the finite set of states represents user-defined context profiles and the HMM models the transition processes through the states, or rather the behavior of a user transitioning from one context to another.

Applications that have used HMMs in learning, classifying and modeling the dynamics of different situations

include tracking the daily activity of residents in an assisted living community [6]. Using wearable sensors to generate personal contextual annotations from an audio-visual recording of a meeting [32]. Inferring environmental context by recognizing classes of sound [33]. Recognizing a person's situation from a wearable audio-visual system [34]. Identifying bathroom activities based on sound event classification [35]. And finally an application that automatically tracks the progress of maintenance or assembly tasks using body worn sensors [36].

Just as BNs, HMMs also require a training phase in order to classify activities. Optimally this learning phase should occur without intervention from the user, especially to be applicable in Body Sensor Networks. Clarkson and Pentland in [33] propose a method using incremental additive learning to avoid relying on the user for learning. As such, if the learned system is given an event that cannot be recognized accurately by the current HMM, the system can recognize this by the indication of low scores and then generates a new model for the event.

When modeling sequences of events, it is advantageous that HMMs allow for time variance (the event may be performed at varying speeds), repetition (an event may be repeated any number of times) [36] and they are able to deal with variable length sequences. In addition, due to their probabilistic framework they are able to take into account noisy sensors and imperfect training data coming from different sources of uncertainty. Overall, HMMs build a statistical memory of sequences of events that prove to be relatively robust with regards to temporal changes and allow for high-level domain knowledge to be incorporated into the model [3].

**Limitations:** One disadvantage of HMMs includes not being able to extract a valuable model of event sequences if activity transitions in the data set do not occur often enough [18]. In addition, if the trained model is too general then it will classify a larger set of events with a high probability, thus not being able to distinguish between specific events that may be somewhat similar [33]. HMMs may be rather limited to represent models in the real-world domain because their notion of state lacks the structure that exists in the real-world [24]. Lastly, HMMs can be computationally expensive in regards to performance because the HMM requires enumerating all possible paths through the model [3].

Next, we present a hierarchical structure of HMMs. **Table 5** describes the advantages and limitations of the HMM approach, and **Table 6** shows examples of HMMs used in sensor networks.

#### 4.3.1. Hierarchical Hidden Semi-Markov Models

In context awareness sensing there is a natural hierarchy of human activities, especially as the activities become more and more complex. So it is beneficial to model various sequences of events using HMMs at higher levels

**Table 5. Benefits and drawbacks of the Hidden Markov Model approach to context awareness.**

Technique	Advantages	Limitations
Hidden Markov Model	<ol style="list-style-type: none"> <li>1) Smoothness constraint: allow for continual recognition of context or sequences of activities.</li> <li>2) Allows for time variance (an action may be performed at varying speeds) and repetition (an action may be repeated any number of times).</li> <li>3) Handles noisy sensors and imperfect training data due to their probabilistic framework.</li> </ol>	<ol style="list-style-type: none"> <li>1) Requires a training phase to classify activities, so needs user participation.</li> <li>2) If activity transitions in the dataset don't occur often enough, HMMs might not extract a valuable model of event sequences.</li> <li>3) If trained model is too general, can't distinguish between very similar events.</li> <li>4) Limited representation of models in the real-world domain.</li> <li>5) Computationally expensive in regards to performance.</li> </ol>
Hierarchical Hidden Semi-Markov Model	<ol style="list-style-type: none"> <li>1) Reasons about relative order, duration and abstraction of general activity from sequences of sub-activities.</li> <li>2) Allows for sequences of sub-activities to be recognized without the system having to learn a separate HMM for each sequence.</li> <li>3) Improves computational performance by not processing the training of useless sub-activities.</li> </ol>	<ol style="list-style-type: none"> <li>1) May not scale to environments that contain hundreds of sensors.</li> <li>2) Requires user participation.</li> </ol>

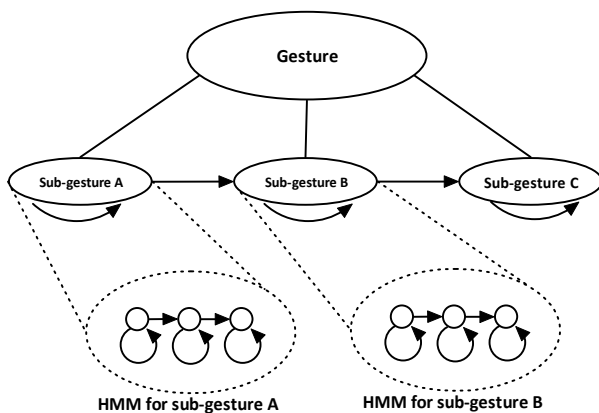
**Table 6. Examples of HMM techniques in sensor networks.**

	Application	Experiment	Results
Kern <i>et al.</i> (2002) [32]	Uses wearable computers and sensor systems to generate personal contextual annotations in audio visual recordings of meetings.	HMMs used to segment an audio stream into sequences assigned to different speakers.	Use of only HMMs results in a recognition error of 18%. Study shows that many issues in recognition of complex postures and gestures remain.
Clarkson <i>et al.</i> (1998) [33]	Uses HMMs as a statistical/pattern recognition framework to recognize environmental context by audio classification.	The HMM classification framework is integrated with Nomadic Radio, a wearable messaging system, to create a wearable platform that is aware of its audio environment.	System proved to be robust towards noise and reliably detected speech by and around the user.
Clarkson <i>et al.</i> (2000) [34]	Presents modeling with HMMs to recognize a person's situation; situations are coarse locations (at work, in a grocery store) and coarse events (having a conversation, walking on the street).	A wearable camera and microphone are used to model 12 dimensions of video features and 12 dimensions of audio features, in order to recognize a person's context.	Results show a surprisingly high accuracy considering the lack of context and the coarse features; however can't assume this method will work for a wide variety of conditions and situations.
Chen <i>et al.</i> (2005) [35]	Uses HMM parameters for accurate and robust recognition and classification of major activities occurring within a bathroom based on sound.	Experiments first performed in a constrained setting and then in an actual trial involving real people using their bathroom in the normal course of their daily lives.	Preliminary results show an accuracy rate of above 84% for most sound categories.
Lukowicz <i>et al.</i> (2004) [36]	Automatically tracks the progress of maintenance or assembly tasks using body worn sensors; uses body worn microphones and accelerometers.	Technique is applied to activities in a wood shop.	On a simulated assembly task, the system can successfully segment and identify most shop activities in a continuous data stream with 84.4% accuracy.
Kautz <i>et al.</i> (2003) [28]	Proposes hierarchical hidden semi-Markov models for tracking the daily activities of residents in an assisted living community.	Implements HHSMM to link location and movement information to a subject's behavior, and reasons about the hierarchical relationship between abstract actions and sub-actions, and both qualitative and quantitative metric constraints.	The semi-Markov structure allows the network to distinguish different activities solely based on their duration. Shows that better algorithms and representations are needed to scale up to larger, more detailed models and fine-grained low-level input data.
Chambers <i>et al.</i> (2002) [37]	Uses an extension of HMM for hierarchical recognition of complex human gestures for sports video annotation.	Uses a hierarchy of HMMs and accelerometers to extract complex gesture recognition; dataset consists of several Kung Fu martial art movements acted out by an instructor in a simulated training video.	The system can robustly differentiate between gestures and accurately segment a gesture into its comprising sub-gestures.

of abstraction through a hierarchy. For example if a person is “having a meal,” this activity consists of a variety of sub-activities such as preparing the food, setting the table, eating the meal, etc. Reasoning about the relative order and duration of the sub-activities, and abstracting the general activity of “having a meal” from the sub-activities can be accomplished with a hierarchy of HMMs. Individual movements would be represented by a number of standard HMMs at the lowest level of the hierarchy. Higher layers represent the combination of movements, and at even higher layers sequences of combination movements can be recognized [37]. As such, the system as a whole is able to model the various levels of complexity of a general activity. The HMM at each level only interacts with the levels directly above and below it; in other words a node at one level may have a transition to a node in the HMM at the direct higher or lower level [28]. Some applications of hierarchical HMMs include recognizing complex human gestures for video annotation (see **Figure 5**) [37] and tracking daily activities [28].

**Advantages of Hierarchical HMMs:** An advantage of the multi-layer HMM is that it allows a rich sequence of sub-activities to be recognized without the system having to learn a separate HMM for each sequence; whereas the standard HMM would require training of a new model with each new combination sequence [37]. This is especially beneficial because the order of sub-activities performed to compose the ultimate activity will differ from person to person, and will differ even when a single person performs the same activity a number of times. Yet the system is still able to recognize the highest level activity. In addition, the activity may be performed with some sub-activities occurring that are not related at all to the general activity. Thus if hierarchical HMMs don't require each new sequence to be learned, the system avoids training useless sub-activities, hence improves the computational performance of the system as a whole.

Layered HMMs (LHMM) are separate layers of HMMs



**Figure 5. Hierarchical hidden Markov model for complex gesture recognition.**

connected according to their inferential results. In this way, each level of the LHMM hierarchy can be trained in a different way. Instead of having a whole HMM responsible for all actions, LHMMs allow for different layers to be responsible for detecting certain actions. There are two ways to do inference with LHMMs. One way is to select the model in which the highest likelihood is selected. This information is made available as input to the HMM at the next level in the hierarchy. In the distributional approach, the full probability distribution over the models is passed to the higher-level HMMs. It may be beneficial for a system to recognize different activities solely based on metric time or the duration the system spends in each sub-step of an activity. This can be achieved by associating a probability distribution over the time the system remains in one state before it transitions to another state [28]. Thus metric, non-exponential time is added to the system resulting in the hierarchical hidden semi-Markov Model.

## 5. Conclusions

Context awareness in Body Sensor Networks allows for the current state of the user and environment to reason about physical and biochemical sensor signals. Context-aware sensing is an integral part of the BSN design in order to allow for long-term pervasive health care monitoring of patients. The outcome of such monitoring would prevent mortality on the grounds that BSNs could have detected precursors of the death.

In this paper, we presented context-aware approaches including Artificial Neural Network (ANN), Bayesian Network (BN) and Hidden Markov Model (HMM). No technique is the “best” for deducing context awareness and each method addresses different issues that arise from context-aware sensing in Body Sensor Networks.

## 6. References

- [1] A. K. Dey, D. Salber, G. D. Abowd and M. Futakawa, “The Conference Assistant: Combining Context-Awareness With Wearable Computing,” *Proceedings of the 3rd International Symposium on Wearable Computers*, San Francisco, 1999, pp. 21-28.
- [2] A. Krause, D. Siewiorek, A. Smailagic and J. Farringdon, “Unsupervised, Dynamic Identification of Physiological and Activity Context in Wearable Computing,” *Proceedings of the 7th International Symposium on Wearable Computers*, White Plains, 2003, pp. 88-97.
- [3] G.-Z. Yang, “Body Sensor Networks,” Springer-Verlag, London, 2006.
- [4] B. Fritzke, “Some Competitive Learning Methods, Artificial Intelligence Institute,” Dresden University of Technology, Dresden, 1997.
- [5] E. Elnahrawy and B. Nath, “Poster Abstract: Online Data

- Cleaning in Wireless Sensor Networks,” *Proceedings of the 1st International conference on Embedded networked sensor systems*, Los Angeles, 2003, pp. 294-295.
- [6] S. Thiemjarus, B. P. L. Lo and G.-Z. Yang, “A Distributed Bayesian Framework for Body Sensor Networks,” *IEE Proceedings of the 2nd International Workshop on Body Sensor Networks*, 2005, pp. 24-26.
- [7] K. van Laerhoven, K. A. Aidoo and S. Lowette, “Real-time Analysis of Data from Many Sensors with Neural Networks,” *Proceedings of the 5th International Symposium on Wearable Computers*, Los Angeles, 2001, pp. 115-123.
- [8] S. Thiemjarus, R. King, B. Lo, D. Gilles and G.-Z. Yang, “A Noise Resilient Distributed Inference Framework for Body Sensor Networks,” *Proceedings of the 3rd International Conference on Pervasive Computing*, Munich, 2005, pp. 13-18.
- [9] P. Nurmi, P. Floreen, M. Przybilski and G. Linden, “A Framework for Distributed Activity Recognition in Ubiquitous Systems,” *Proceedings International Conference on Artificial Intelligence (ICAI)*, Las Vegas, 2005, pp. 650-655.
- [10] S. Thiemjarus, B. P. L. Lo and G.-Z. Yang, “Context Aware Sensing—What is the Significance?” *IEE Proceedings of the Perspective in Pervasive Computing*, London, 2005, pp. 163-170.
- [11] K. van Laerhoven, “Combining the Self-Organizing Map and K-Means Clustering for On-Line Classification of Sensor Data,” *Artificial Neural Networks—ICANN*, Vienna, 2001, pp. 464-470.
- [12] A. Schmidt, K. A. Aidoo, A. Takaluoma, U. Tuomela, K. van Laerhoven and W. van de Velde, “Advanced Interaction in Context,” *Proceedings of the 1st International Symposium on Handheld and Ubiquitous Computing*, Karlsruhe, 1999, pp. 89-101.
- [13] M. Lagerholm, C. Peterson, G. Braccini, L. Edenbrandt and L. Sormmo, “Clustering ECG Complexes Using Hermite Functions and Self-Organizing Maps,” *IEEE Transactions on Biomedical Engineering*, Vol. 47, No. 7, July 2000, pp. 838-848.
- [14] K. van Laerhoven and O. Cakmake, “What Shall We Teach Our Pants?” *Proceedings of the 4th International Symposium on Wearable Computers*, Atlanta, 2000, pp. 77-83.
- [15] K. Simelius, M. Stenroos, L. Reinhardt, J. Nenonen, I. Tierala, M. Makijarvi, L. Toivonen and T. Katila, “Spatiotemporal Characterization of Paced Cardiac Activation with Body Surface Potential Mapping and Self-Organizing Maps,” *Physiological Measurements*, Vol. 24, No. 3, 2003, pp. 805-816.
- [16] D. Gao, M. Madden, M. Schukat, D. Chambers and G. Lyons, “Arrhythmia Identification from ECG Signals with a Neural Network Classifier Based on a Bayesian Framework,” *Proceedings of the 24th SGAI International Conference on Innovative Techniques and Applications of Artificial Intelligence*, Cambridge, 2004.
- [17] S. Thiemjarus, B. P. L. Lo and G.-Z. Yang, “A Spatio-Temporal Architecture for Context-Aware Sensing,” *Proceedings of the IEEE International Workshop on Wearable and Implantable Body Sensor Networks*, Cambridge, 2006, pp. 191-194.
- [18] K. van Laerhoven, N. Kern, H. W. Gellersen and B. Schiele, “Towards a Wearable Inertial Sensor Network,” *Proceedings of IEE Eurowearable*, Stevenage, 2003, pp. 125-130.
- [19] M. Philipose, K. Fishkin, M. Perkowitz, D. Fox, H. Kautz and D. Patterson, “Inferring Activities from Interactions with Objects,” *IEEE Pervasive Computing*, Vienna, 2004, pp. 50-57.
- [20] R. Picard, E. Vyzas and J. Healey, “Towards Machine Emotional Intelligence: Analysis of Affective Physiological State,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2001, pp. 1175-1191.
- [21] K. Noguchi, P. Somwong, T. Matsubara and Y. Nakauchi, “Human Intention Detection and Activity Support System for Ubiquitous Autonomy,” *Proceedings of IEEE International Symposium on Computational Intelligence in Robotics and Automation*, 2003, pp. 906-911.
- [22] T. Barger, D. Brown and M. Alwan, “Health-Status Monitoring through Analysis of Behavioral Patterns,” *IEEE Transactions on Systems, Man and Cybernetics: Part A*, Edmonton, 2005, pp. 22-27.
- [23] A. Madabhushi and J. K. Aggarwal, “A Bayesian Approach to Human Activity Recognition,” *Proceedings of the 2nd IEEE International Workshop on Visual Surveillance*, Fort Collins, 1999, pp. 25-32.
- [24] P. Korpipaa, M. Koskinen, J. Peltola, S. M. Makela and T. Seppanen, “Bayesian Approach to Sensor-Based Context Awareness,” *Personal and Ubiquitous Computing*, Vol. 7, No. 2, 2003, pp. 113-124.
- [25] E. M. Tapia, S. S. Intille and K. Larson, “Activity Recognition in the Home Using Simple and Ubiquitous Sensors,” *Proceedings on the 2nd International Conference on Pervasive Computing*, Vienna, 2004, pp. 158-175.
- [26] E. Elnahrawy and B. Nath, “Context-Aware Sensors,” *Proceedings on 1st European Workshop on Wireless Sensor Networks*, Berlin, 2004, pp. 73-93.
- [27] P. Domingos and M. Pazzani, “Beyond Independence: Conditions for the Optimality of the Simple Bayesian Classifier,” *Proceedings of the 13th International Conference on Machine Learning*, Bari, 1996, pp. 105-112.
- [28] H. Kautz, O. Etzioni, D. Fox and D. Weld, “Foundations of Assisted Cognition Systems,” Technical report CSE-02-AC-01, University of Washington, Department of Computer Science and Engineering, Washington, D.C., 2003.
- [29] M. Angermann, P. Robertson and T. Strang, “Issues and Requirements for Bayesian Approaches in Context Aware Systems,” *Location- and Context Awareness: First International Workshop*, Vol. 3479, 2005, pp. 235-243.
- [30] R. King, L. Atallah, A. Darzi and G.-Z. Yang, “A HMM Framework for Optimal Sensor Selection with Applications to BSN Sensor Glove Design,” *Proceedings of the Fourth Workshop on Embedded Networked Sensors*, Cork,

- 2007, pp. 58-62.
- [31] N. Oliver, E. Horvitz and A. Garg, "Layered Representations for Human Activity Recognition," *Proceedings of the 4th IEEE International Conference on Multimodal Interfaces*, Pittsburgh, 2002, pp. 3-8.
- [32] N. Kern, B. Schiele, H. Junker and G. Troster, "Wearable Sensing to Annotate Meeting Recordings," *Proceedings on the 6th International Symposium on Wearable Computers*, Seattle, 2002, pp. 186-194.
- [33] B. Clarkson and A. Pentland, "Extracting Context from Environmental Audio," *Proceedings of the 2nd International Symposium on Wearable Computers*, Pittsburgh, 1998, pp. 154-155.
- [34] B. Clarkson, K. Mase and A. Pentland, "Recognizing User Context via Wearable Sensors," *Proceedings of the 4th International Symposium on Wearable Computers*, Atlanta, 2000, pp. 69-76.
- [35] J. Chen, A. H. Kam, J. Zhang, N. Liu and L. Shue, "Bathroom Activity Monitoring Based on Sound," *Proceedings of the 3rd International Conference on Pervasive Computing*, Munich, 2005, pp. 47-61.
- [36] P. Lukowicz, J. A. Ward, H. Junker, M. Stager, G. Troster, A. Atrash and T. Starner, "Recognizing Workshop Activity Using Body Worn Microphones and Accelerometers," *Pervasive Computing*, Vol. 3001, 2004, pp. 18-32.
- [37] G. S. Chambers, S. Venkatesh, G. A. W. West and H. H. Bui, "Hierarchical Recognition of Intentional Human Gestures for Sports Video Annotation," *Proceedings of the 16th IEEE Conference on Pattern Recognition*, Québec City, Vol. 2, 2002, pp. 1082-1085.
- [38] B. T. Korel and S. G. M. Koo, "Addressing Context Awareness Techniques in Body Sensor Networks," *Proceedings of the 21st IEEE International Conference on Advanced Information Networking and Applications (AINA 2007)*, Niagara Falls, Vol. 2, 2007, pp. 798-803.
- [39] G. D. Abowd and A. K. Dey, "Towards a Better Understanding of Context and Context-Awareness," Georgia Institute of Technology, Technical Report GIT-GVU-99-22, Atlanta, 1999.
- [40] G. D. Abowd, A. K. Dey, R. Orr and J. Brotherton, "Context-Awareness in Wearable and Ubiquitous Computing," *Proceedings of the 1st International Symposium on Wearable Computers*, Cambridge, 1997, pp. 179-180.
- [41] U. Anliker, H. Junker, P. Lukowicz and G. Troster, "Design Methodology for Context-Aware Wearable Sensor System," *Pervasive Computing: 3rd International Conference*, Seattle, 2005, pp. 220-236.
- [42] A. R. Golding and N. Lesh, "Indoor Navigation Using a Diverse Set of Cheap, Wearable Sensors," *Proceedings of the 3rd International Symposium on Wearable Computers*, San Francisco, 1999, pp. 29-36.
- [43] J. Himberg, J. A. Flanagan and J. Mantyjarvi, "Towards Context Awareness Using Symbol Clustering Map," *Proceedings of Workshop for Self-Organizing Maps*, Kitakyushu, 2003, pp. 249-254.
- [44] N. Kern and B. Schiele, "Context-Aware Notification for Wearable Computing," *Proceedings of the 7th International Symposium on Wearable Computers*, Osaka, 2003, pp. 223-230.
- [45] P. Langley, W. Iba and K. Thompson, "An Analysis of Bayesian Classifiers," *Proceedings of the 10th National Conference on Artificial Intelligence*, San Jose, 1992, pp. 223-228.
- [46] B. P. L. Lo, S. Thiemjarus, R. King and G.-Z. Yang, "Body Sensor Network—A Wireless Sensor Platform for Pervasive Healthcare Monitoring," *Proceedings of the 3rd International Conference on Pervasive Computing*, Munich, 2005, pp. 77-80.
- [47] B. P. L. Lo and G.-Z. Yang, "Key Technical Challenges and Current Implementations of Body Sensor Networks," *IEE Proceedings of the 2nd International Workshop on Body Sensor Networks*, London, 2005, pp. 1-5.
- [48] D. Macq, M. Verleysen, P. Jaspers and J. D. Legat, "Analog Implementation of a Kohonen Map with On-Chip Learning," *IEEE Transactions on Neural Networks*, 1993, pp. 456-461.
- [49] D. Patterson, D. Fox, H. Kautz and M. Philipose, "Expressive, Tractable and Scalable Techniques for Modeling Activities of Daily Living," *Proceedings of the 2nd International Workshop on Ubiquitous Computing for Pervasive Healthcare Applications*, 2003.
- [50] N. Ravi, N. Dandekar, P. Mysore and M. L. Littman, "Activity Recognition from Accelerometer Data," *Proceedings of IAAI*, Pittsburgh, 2005.
- [51] K. van Laerhoven, "Teaching Contexts to Applications," *Proceedings of the Workshop on Situated Interaction in Ubiquitous Computing*, 2000, pp. 5-8.
- [52] K. van Laerhoven, B. P. L. Lo, J. W. P. Ng, S. Thiemjarus, R. King, S. Kwan, H. W. Gellersen, M. Sloman, O. Wells, P. Needham, N. Peters, A. Darzi, C. Toumazou and G.-Z. Yang, "Medical Healthcare Monitoring with Wearable and Implantable Sensors," *International Workshop on Ubiquitous Computing for Pervasive Healthcare Applications*, 2004.
- [53] K. van Laerhoven, A. Schmidt and H. W. Gellersen, "Multi-Sensor Context Aware Clothing," *Proceedings of the 6th International Symposium on Wearable Computers*, Seattle, 2002, pp. 49-57.