Posture Detection with Body Area Networks *

Ioannis Ch. Paschalidis†
yannisp@bu.edu

Wuyang Dai‡
wydai@bu.edu

Dong Guo
dong.dongguo@gmail.com

Yingwei Lin
yingwei@bu.edu

Keyong Li
likeyong@ieee.org

Binbin Li
libinbin@bu.edu

ABSTRACT

Body posture detection is extremely useful in health monitoring and rehabilitation. We develop a method to detect body posture that uses signal strength measurements from sensor nodes forming a Wireless Body Area Network (WBAN). We assume that postures (formations) take values in a discrete set and develop a composite hypothesis testing approach which uses a Generalized Likelihood Test (GLT) decision rule. The GLT rule distinguishes between a set of probability density function (pdf) families constructed using a custom pdf interpolation technique. The GLT is compared with the simple Likelihood Test (LT). We also adapt one prevalent supervised learning approach, Multiple Support Vector Machine (MSVM), to compare with our probabilistic methods. Due to the highly variant measurements from the WBAN, and these methods' different adaptability to multiple observations, our analysis and experimental results suggest that GLT is more accurate and suitable for posture/formation detection. Even for very similar postures in our experiments, GLT demonstrates high detection accuracy (around 97% with multiple observations). Besides the body area networks, the formation detection problem has interesting applications in autonomous robot systems.

Categories and Subject Descriptors

I.5 [**Pattern Recognition**]: Miscellaneous

*Research partially supported by the NSF under grant EFRI-0735974, by the DOE under grant DE-FG52-06NA27490, by the ARO under grant W911NF-11-1-0227, by the ODDR&E MURI10 program under grant N00014-10-1-0952, and by a Dean's catalyst award from the College of Engineering at Boston University.

†Corresponding author. Dept. of Electrical & Computer Eng., and Division of Systems Eng., Boston University, 8 Saint Mary's St., Boston, MA 02215.

‡All authors are with Center for Information & Systems Eng., Boston University

BODYNETS 2011, November 07-08, Beijing, People's Republic of China

DOI 10.4108/icst.bodynets.2011.247212

General Terms

Algorithm, Performance, Theory

Keywords

Posture detection, Generalized likelihood test, Multiple observation classification.

1. INTRODUCTION

Wireless Body Area Networks (WBANs) consist of small, battery powered, wireless sensor nodes attached (or implanted) to the human body. Interesting sensors include pacemakers, implantable cardioverter-defibrillators (ICDs), pulse oximeters, glucose level sensors, sensors to monitor the heart (ECG, blood pressure, etc.), thermometers, or sensors to monitor body posture. Several devices incorporating such sensing capabilities and wireless connectivity exist today (at least as prototypes) [12, 11]. The emergence of WBANs could potentially revolutionize health monitoring and health care, allowing for instance remote uninterrupted monitoring of patients in their normal living environment. Broad applications are envisioned in medicine, military, security, and work place [9, 17].

One particular service that is of great interest in many of these application contexts is to detect body posture. We give a few examples. Work related back injuries are a frequent source of litigation. Many such injuries can be avoided if an incorrect posture for lifting heavy packages can be detected early on and corrected by proper training. In another example, the inhabitants of a smart home will be able to control different functions (such as heating or cooling, lighting, etc.) merely by gesture. Such functionality can be a blessing for the elderly and the disabled. An additional application involves monitoring inhabitants of assisted living facilities; body posture reveals a lot about the "state" of an individual and could alert caregivers in case of emergency (e.g., body falling to or lying on the stairs [10]).

The basic idea is to place WBAN devices on different parts of the body, say, wrist, ankle, shoulder, knee, etc., and to detect the posture of the body through the *formation* of the wireless sensor nodes, i.e., the relative positions of these nodes. The premise of this paper is that the formation of the wireless sensor nodes is reflected by the Received Signal Strength Indicators (RSSI) at WBAN nodes. As our experiments show, this relationship indeed holds, but in a rather complicated way. In particular, the RSSI signatures of a formation (or posture) do not have a simple dependence on the pairwise distances of the WBAN nodes. Instead, they

are correlated among different WBAN pairs, do not follow a standard noise model, and they also depend on the time and the location of the body and also other subtle aspects (e.g., the thickness of clothes). This is the main reason we focus on measurement-based methods, including probabilistic classifiers and supervised learning approaches.

The problem at hand has wider applicability than the WBAN setting. The techniques we develop are also applicable in detecting (and controlling) the formation of robot swarms deployed in the interior of a building. In indoor deployments, the mapping from RSSI to distance is erratic and unpredictable, requiring the more sophisticated classification or hypothesis testing techniques we develop (see also [16, 13]).

In this work, we assume that formations take values in a discrete set and formulate the formation detection problem as a composite hypothesis testing problem. We use the *Generalized Likelihood Test (GLT)* between a set of pdf families characterizing each possible formation (posture). These pdf families are constructed using a generalized joint pdf interpolation scheme; a simpler interpolation scheme was proposed by [3]. Comparisons with a simple *Likelihood Test (LT)* and a supervised learning approach, the *Multiple Support Vector Machine (MSVM)* [4, 5], are addressed by our simulation and experimental results.

For simulations and experiments, we conduct the testing for both a single observation (RSSI measurements at a certain time) and multiple observations (a sequence of measurements). The results show that GLT has several potential advantages comparing to MSVM:

- it results in better handling of multiple observations;
- is robust to measurement uncertainty;
- and is computationally more efficient for multi-class classification.

The rest of the paper is organized as follows: In Section 2, we introduce the recent research on posture detection in WBANs and also introduce the GLT and SVM methods with some other applications. In Section 3, we define the problem. In Section 4, we introduce the decision rules for the hypothesis testing formulation. In Section 5, we discuss the MSVM approach. In Section 6 and Section 7, we describe simulation and experimental results and provide analysis, respectively. Conclusions are in Section 8.

Notation: We use bold lower case letters for vectors and bold upper case letters for matrices. All vectors are column vectors and we write $\mathbf{x} = (x_1, \cdots, x_n)$ for economy of space. Transpose is denoted by prime.

2. RELATED WORK

Recent developments in sensor technology make wearable sensors and the resulting body area network applicable to a variety of scenarios. Body posture detection in particular, has been studied for different purposes and with different approaches. One application concerns detecting a fall of the monitored individual: early falling detection [10], which is useful in protecting senior citizens. Although video monitoring or alarms with buttons could offer alternatives, they have their own limitations. The former one raises privacy concerns while the latter one requires the senior person's consciousness after falling. A WBAN solution does not suffer from these limitations [10].

A custom-designed WBAN for posture detection is constructed in [7]. The constructed WBAN uses accelerometers to acquire information about the posture, and then makes a classification according to a pre-set table. The accelerometers are also used in [8] for posture detection in ambulatory monitoring. As shown in these previous works, it is common to use accelerometers as the main source of relevant data ([10, 7, 8]). Accelerometers indeed provide accurate measurements for quick movements and thus are more suitable for motion detection. However, for posture detection, they need an inference step to "derive" posture from motion, which makes the detection more complicated and highly relying on the logical rules in this inference step.

A novel approach for posture detection by using the relative proximity information (based on RSSI) between sensor nodes is proposed in [15]. Compared to accelerometer-based methods, this approach is not limited to activity intensive postures such as walking and running, but also works for low activity postures such as sitting and standing [15]. This is one of the main reasons we choose RSSI signals for posture detection. Different from the Hidden Markov Model (HMM) used in [15], we ignore temporal relations among measurements to make our algorithm more robust. Using discrete multiple postures, we formulate posture detection as a composite hypothesis testing or a multi-class classification problem.

For classification problems, SVM is a well known classifier for its generally good performance in many applications of binary classification [2]. To fit SVM from binary class classification to multiple class classification, it is common to apply it between each pair of classes and then do a simple voting to make a final decision [5]. We call this extension Multiple SVM (MSVM). Besides MSVM, there are probabilistic approaches for multiple class classification, where the classification is viewed as a composite hypothesis testing problem. Although, generally, probabilistic approaches (e.g., GLT) are not as good as statistical approaches (e.g., SVM) due to the difficulty of precise density estimation, under certain conditions, GLT provides an asymptotically best solution [18]. It is worth noting that a recent study of GLT shows a good performance in an application of multiple class classification [13], and also demonstrate its ability to accommodate noise.

In our own work, the analysis and experiments indicate GLT's suitability and that it merits further investigation in the future.

3. FORMULATION

Consider k sensors, where one of them is the receiver and the rest are transmitters, and let $\mathcal{C} = \{1, \ldots, C\}$ be a discrete set of their possible formations. In practice, the positions of the sensors take values in a continuous space and one can argue that formations are also continuous. However, for many applications, including the ones discussed in the Introduction, we are interested in distinguishing between relatively few formations which characterize the "state" of the underlying physical system (the body, the robot swarm).

The discretization of formations is in line with our earlier sensor localization work [16, 13, 14]. It makes the detection/classification problem more tractable but introduces the requirement that the techniques to be used should be robust enough and tolerant to mild or moderate perturbations. Namely, every element of $\mathcal{C}$ represents a "family" of similar

looking formations that can be generated from a nominal formation subject to perturbations.

The RSSI measurements are denoted by a column vector $\mathbf{y} \in \mathbb{R}^d$, where $d = k - 1$. In each of the methods we will present, the formation classifier is computed from a *training* set of RSSI measurements, and then we examine experimentally how well the classifiers generalize to additional measurements. Two types of methods for building the classifier will be considered next, including a probabilistic hypothesis testing approach and MSVM. We first use simulation data to analyze the properties of our proposed method and then test the performance on real sensor data.

4. PROBABILISTIC APPROACH

In the probabilistic approach, we treat each formation as a composite hypothesis associated with a family of pdfs in the space of the joint RSSI measurements. We use a family of pdfs for each formation in order to improve system robustness (e.g., with respect to time and location). The pdfs are first estimated from the training data, employing a technique combining a Parzen windowing scheme and Gaussianization [6]. The pdf families are formed using a pdf interpolation technique that we have generalized from [3]. Finally, decisions are made according to the well-known (generalized) maximum likelihood test.

4.1 Multivariate density estimation

Suppose among the M samples, $\mathbf{y}_1, \ldots, \mathbf{y}_m$ are associated with one formation in $\mathcal{C}$. Let $\mathbf{Y} = [\mathbf{y}_1 \mathbf{y}_2 \cdots \mathbf{y}_m]$. We view the measurements $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_m$ as realizations of a random variable $\mathbf{y} = (y_1, y_2, \ldots, y_d)$. We first estimate the marginal pdfs of $\mathbf{y}$ denoted by $p_i(y_i)$, $i = 1, \ldots, d$, using Parzen windows (see Appendix A). The benefit of using Parzen windows is that the resulting $p_i(y_i)$'s are smoothened.

We then estimate the joint pdf using the Gaussianization method of [6], the basic assumption (or approximation) of which is : when we transform the marginal distributions separately into Gaussian distributions, the joint distribution also becomes Gaussian. Specifically, we construct an element-wise Gaussianization function $\mathbf{h(y)} = (h_1(y_1), h_2(y_2), \ldots, h_d(y_d))$, such that the marginal distributions of $\mathbf{z = h(y)}$ are zero-mean Gaussian distributions. Then, the method of [6] assumes $\mathbf{z}$ is also jointly Gaussian, thus, its pdf can be determined from the sample covariance matrix $\mathbf{\Sigma_z}$. Then, the joint pdf of $\mathbf{y}$ can be estimated as (see Appendix B).

$$p(\mathbf{y}) = \frac{\mathbf{g_{\Sigma_z}(h(y))}}{|\nabla \mathbf{h^{-1}(h(y))}|} = \mathbf{g_{\Sigma_z}(h(y))} \prod_{i=1}^{d} \frac{p_i(y_i)}{g_1(h_i(y_i))}, \quad (1)$$

where $\mathbf{g_{\Sigma_z}}$ denotes a zero-mean multivariate Gaussian density function with covariance $\mathbf{\Sigma_z}$ and g_1 denotes a zero-mean univariate Gaussian density function with unit variance.

4.2 Interpolation of probability density functions

In order to construct a family of pdfs for each formation, we introduce an interpolation technique for probability density functions.

a) Interpolation formula: Let each $f_i(\mathbf{y})$, $i = 1, \ldots, N$, be a d-dimensional pdf with mean $\boldsymbol{\mu}_i$ and covariance matrix $\mathbf{K}_i$. Note that these are generally non-Gaussian pdfs. We call what follows the *linear interpolation* of these pdfs with a weight vector $\boldsymbol{\alpha}$, where the elements of $\boldsymbol{\alpha}$ are nonnegative and sum to one.

It is desirable that the mean and covariance of the interpolated pdf equal

$$\boldsymbol{\mu} = \sum_{i=1}^{N} \alpha_i \boldsymbol{\mu}_i, \qquad \mathbf{K} = \sum_{i=1}^{N} \alpha_i \mathbf{K}_i. \quad (2)$$

Define a coordinate transformation for each $i = 1, \ldots, N$, so that $\mathbf{y}_i$ is defined by

$$\mathbf{K}^{-1/2}(\boldsymbol{y} - \boldsymbol{\mu}) = \mathbf{K}_i^{-1/2}(\mathbf{y}_i - \boldsymbol{\mu}_i), \quad (3)$$

where $\mathbf{K}^{1/2}(\mathbf{K}^{1/2})' = \mathbf{K}$. The Jacobian of each transformation is expressed as

$$J_i = \sqrt{\det(\mathbf{K}_i \mathbf{K}^{-1})}. \quad (4)$$

The interpolation formula is then

$$f_{\boldsymbol{\alpha}}(\mathbf{y}) = \sum_{i=1}^{N} \alpha_i J_i f_i(\mathbf{y}_i). \quad (5)$$

This interpolation not only achieves property (2), but also preserves the "shape" information of the original pdfs to a large extent. For example, if the original pdfs are Gaussian, then the interpolated pdf is also Gaussian. This cannot be achieved by, say, a simple weighted sum of the original pdfs. The formula above was first given in [3], but formally only for cases satisfying $d = N$. We next verify that the general case is also true.

b) Verification: We first check that it is a legitimate probability measure:

$$\begin{aligned}
&\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{\boldsymbol{\alpha}}(\mathbf{y}) dy_1 \cdots dy_d \\
&= \sum_{i=1}^{N} \alpha_i \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} J_i f_i(\mathbf{y}_i) dy_1 \cdots dy_d \\
&= \sum_{i=1}^{N} \alpha_i \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_i(\mathbf{y}_i) dy_{i1} \cdots dy_{id} \\
&= \sum_{i=1}^{N} \alpha_i = 1.
\end{aligned}$$

Then we check the mean:

$$\begin{aligned}
&\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \mathbf{y} f_{\boldsymbol{\alpha}}(\mathbf{y}) dy_1 \cdots dy_d \\
&= \sum_{i=1}^{N} \alpha_i \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \mathbf{y} J_i f_i(\mathbf{y}_i) dy_1 \cdots dy_d \\
&= \sum_{i=1}^{N} \alpha_i \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} (\mathbf{K}^{1/2} \mathbf{K}_i^{-1/2} (\mathbf{y}_i - \boldsymbol{\mu}_i) + \boldsymbol{\mu}) \\
&\qquad\qquad f_i(\mathbf{y}_i) dy_{i1} \cdots dy_{id} \\
&= \sum_{i=1}^{N} \alpha_i (\mathbf{K}^{1/2} \mathbf{K}_i^{-1/2} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} (\mathbf{y}_i - \boldsymbol{\mu}_i) f_i(\mathbf{y}_i) \\
&\quad dy_{i1} \cdots dy_{id} + \boldsymbol{\mu} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_i(\mathbf{y}_i) dy_{i1} \cdots dy_{id}) \\
&= \boldsymbol{\mu} \sum_{i=1}^{N} \alpha_i = \boldsymbol{\mu}.
\end{aligned}$$

Lastly, we check the covariance matrix:

$$\begin{aligned}
&\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\mathbf{y}\mathbf{y}'f_{\boldsymbol{\alpha}}(\mathbf{y})dy_1\cdots dy_d\\
=&\sum_{i=1}^{N}\alpha_i\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\mathbf{y}\mathbf{y}'J_if_i(\mathbf{y}_i)dy_1\cdots dy_d\\
=&\sum_{i=1}^{N}\alpha_i\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}(\mathbf{K}^{1/2}\mathbf{K}_i^{-1/2}(\mathbf{y}_i-\boldsymbol{\mu}_i)+\boldsymbol{\mu})\\
&(\mathbf{K}^{1/2}\mathbf{K}_i^{-1/2}(\mathbf{y}_i-\boldsymbol{\mu}_i)+\boldsymbol{\mu})'f_i(\mathbf{y}_i)dy_{i1}\cdots dy_{id}\\
=&\sum_{i=1}^{N}\alpha_i(\mathbf{K}^{1/2}\mathbf{K}_i^{-1/2}\mathbf{K}_i(\mathbf{K}_i^{-1/2})'(\mathbf{K}^{1/2})'+\boldsymbol{\mu}\boldsymbol{\mu}')\\
=&(\boldsymbol{\mu}\boldsymbol{\mu}'+\mathbf{K})\sum_{i=1}^{N}\alpha_i=\boldsymbol{\mu}\boldsymbol{\mu}'+\mathbf{K}.
\end{aligned}$$

Our verification is complete.

4.3 LT and GLT

We associate a hypothesis H_j to each formation $j\in\mathcal{C}$. For each formation j, we collect measurements from different deployments of the nodes according to j in different environments (e.g., rooms of a building). For each set of measurements, we construct a pdf $f(\mathbf{y}|H_j)$ as outlined in Section 4.1. We interpolate as in Section 4.2 the pdfs corresponding to different deployments of formation j to end up with a pdf family $f_{\boldsymbol{\alpha}}(\mathbf{y}|H_j)$ characterizing this formation. As explained earlier, the key motivation for constructing pdf families is to gain in robustness with respect to perturbations that would naturally arise in any deployment of a formation.

The maximum likelihood test (LT) is based on just a single pdf $f(\mathbf{y}|H_j)$ characterizing formation j. Using n observations (sets of RSSI measurements) $\mathbf{y}_1,\ldots,\mathbf{y}_n$, it identifies formation H_L if

$$L=\arg\max_{j\in\mathcal{C}}\prod_{i=1}^{n}f(\mathbf{y}_i|H_j). \tag{6}$$

The test we propose is a composite hypothesis test using the pdf families $f_{\boldsymbol{\alpha}}(\mathbf{y}|H_j)$. It uses the generalized likelihood test (GLT) which was shown to have desirable optimality properties in [13]. Specifically, it identifies formation H_L if

$$L=\arg\max_{j\in\mathcal{C}}\max_{\boldsymbol{\alpha}}\prod_{i=1}^{n}f_{\boldsymbol{\alpha}}(\mathbf{y}_i|H_j). \tag{7}$$

Our experiments show that GLT outperforms LT.

5. MULTIPLE SUPPORT VECTOR MACHINE

In this section we describe a classification approach using a *Support Vector Machine (SVM)*. An SVM is an excellent two-category classifier [4]. We work with one pair of formations, l_1 and l_2, at a time. To find the support vectors, we solve the following dual form of the soft margin problem (see [4]):

$$\begin{aligned}
\max\quad & -\frac{1}{2}\sum_{i=1}^{M_1}\sum_{j=1}^{M_1}\alpha_i\alpha_jI_iI_jK(\mathbf{y}_i,\mathbf{y}_j)+\sum_{i=1}^{M_1}\alpha_i,\\
\text{s.t.}\quad & \sum_{i=1}^{M_1}\alpha_iI_i=0,\\
& 0\le\alpha_i\le C,
\end{aligned} \tag{8}$$

where $\mathbf{y}_i$'s are the original measurements, $K(\cdot,\cdot)$ is the kernel function, $I_i=\pm1$ is the label of sample i with 1 meaning formation l_1 and -1 meaning formation l_2, and M_1 is the number of samples associated with each formation. Given a measurement $\mathbf{y}$, the SVM categorizes it by computing

$$I_{l_1l_2}(\mathbf{y})=\text{sign}(\sum_{i=1}^{M_1}I_i\alpha_iK(\mathbf{y},\mathbf{y}_i)), \tag{9}$$

where $I_{l_1l_2}(\mathbf{y})$ denotes the output label. Again, 1 means formation l_1 and -1 means formation l_2.

We use one of the commonly used kernel function - Gaussian radial basis function:

$$K(\mathbf{y}_1,\mathbf{y}_2)=\exp(-\frac{\|\mathbf{y}_1-\mathbf{y}_2\|^2}{2\sigma^2}). \tag{10}$$

For a C-class SVM, as in our case, we can apply $C(C-1)/2$ pairwise SVMs, and choose the majority vote as the final decision [5]:

$$L=\arg\max_{i\in\mathcal{C}}\sum_{j\neq i}I_{ij}(\mathbf{y}). \tag{11}$$

Formula (11) is for a single observation classification. With multiple observations, we need another level of majority voting over n observations.

At this point, we could find that for the multi-class classification problem, MSVM needs to run SVM several times to classify a given piece of test data and each run involves more than one class of training data. On the other hand for GLT (or LT), the calculation of the likelihood of test data for a certain hypothesis only needs the training data of that class. This results in a substantial computational efficiency of the latter methods compared to MSVM.

6. SIMULATIONS

6.1 Setup

We simulate two Gaussian distributed classes, and means $(0,0)$ (class 1) and $(1,0)$ (class 2) respectively, covariance matrix $\mathbf{I}_{2\times2}$, where $\mathbf{I}_{2\times2}$ is the identity matrix of rank 2. A sample distribution of the two classes is shown in Figure 1. With one simulation run, we generate 100 training data points and 500 test data points per class. The LT and SVM algorithms are directly applied to these data. For GLT, the training data are randomly split into two subsets with equal number of observations and each subset forms a pdf, then interpolation between these two pdfs are executed as described in Section 4.2 to form two pdf families. After that, the GLT algorithm is carried out. The whole process is repeated 100 times and the average success rates are reported to eliminate the variation of randomly generated data, where the success rate is the fraction of correctly classified test data.

In the above setting, the means of these two classes are fixed. We set up another simulation with uncertain means.

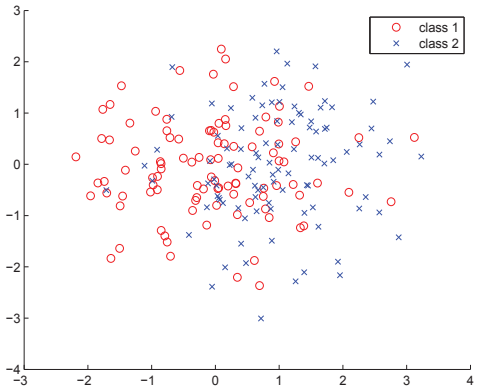

Figure 1: Samples of simulated Gaussian distributed classes.

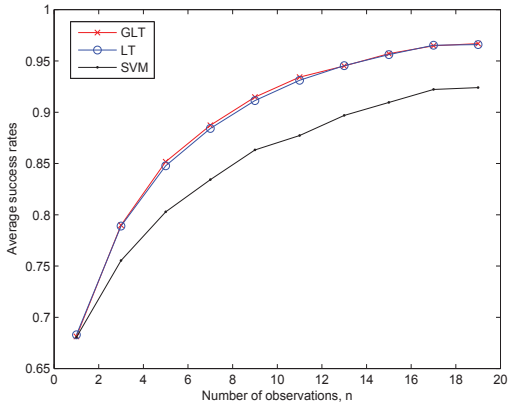

Figure 2: Average success rates of different methods on simulated data.

Noise is added into one dimension of each mean; in particular, the mean of first class is uniformly distributed in the range $(0, 0) \sim (-5, 0)$, while the mean of the second class is uniformly distributed in the range $(1, 0) \sim (1, 5)$. Two training data sets are generated under the extreme values of the means, while the test data are generated under random mean values. The reason that we could choose extreme means for training is that in practice, we ought to gather several sets (much more than just two) of data and extreme situations are likely to be among them.

6.2 Simulation Results

The simulation results with fixed mean values are displayed in Figure 2. The results show that even though all three methods performs equally well for single observation classification, with multiple observations (as described in Section 4.3), probabilistic methods (LT, GLT) have higher success rates.

The results under uncertain means are depicted in Figure 3, which indicates that the GLT is more resistant to uncertainty than LT. SVM is again substantially inferior to GLT.

6.3 Analysis

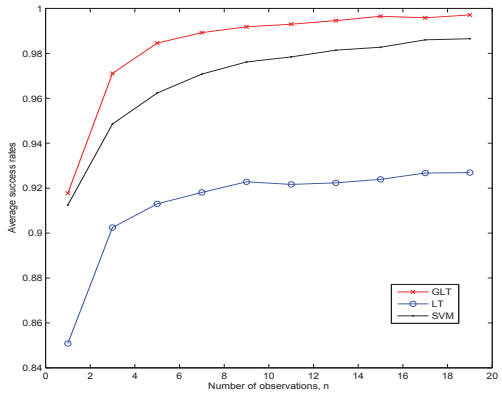

Figure 3: Average success rates of different methods on simulated data with uncertain means.

Our simulation results show that GLT and LT are better for multi-observation test/classification. The intuition behind this is that in expressions (6) and (7), the likelihoods of different observations are multiplied together so that one large likelihood (corresponding to high confidence) can dominate others. This will benefit the test with multiple observations. On the other hand, in MSVM, one observation (no matter how confident) simply add one vote to a class which doesn't make full use of the observations.

For the simulation with uncertain mean values, GLT outperforms LT because GLT has the ability to appropriately shift the density to fit the test data by making use of the interpolated pdf families. This is a unique characteristic of GLT and it results in GLT's robustness with respect to system parameters (i.e., the mean values in this case). It is worth noting that this characteristic of GLT requires the availability of the "extreme" distributions in the training data and is also affected by the mechanism of pdf interpolation. The fact that LT performs even worse than MSVM suggests that modeling each class into one single density distribution does not fit the underlying property of the data.

7. EXPERIMENTS

7.1 Hardware

For our experiments we used the Intel Imote2 motes from Crossbow Technology Inc. The Imote2 (2400-2483.5 MHz band) uses the Chipcon CC2420, IEEE 802.15.4 compliant, ZigBee-ready radio frequency transceiver integrated with an PXA271 micro-controller. Its radio can be tuned within the IEEE 802.15.4 channels, numbered from 11 (2.405 GHz) to 26 (2.480 GHz), each separated by 5 MHz. The RF transmission power is programmable from 0 dBm (1 mW) to -25 dBm.

In order to reduce the signal variation for certain posture, we tuned the RF transmission power to -25 dBm at channel 11.

7.2 Setup

We use 4 sensors attached to the right up chest, outside of left wrist, left pocket and left ankle. It is easy and convenient

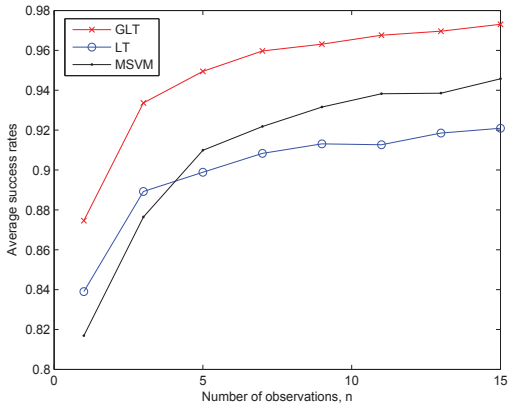

Figure 4: Average success rates of different methods on real sensor data.

to attach sensors at these 4 body areas. Among all these sensors, the right up chest one is used as the receiver while the rest are transmitters.

If the postures are very different (such as standing vs. bending forward), all methods (GLT, LT and MSVM) show very high accuracy which makes the comparison difficult. So, we use three patterns in our experiment, which are not quite easy to differentiate:

- standing straight with hands aligned with the body (military standing at attention);
- standing straight with two hand loosely held together in front of the body;
- and standing straight with two arms folded in front of the chest.

We take three sets of data at different times. Each set contains all postures. Each of the three sets has roughly 1000 observations per posture. In each experiment, we randomly select 200 samples per posture from these three data sets and samples from two of them are used for training, while samples from the rest set are used for testing. We repeat the experiment 60 times and report the average success rates.

7.3 Performance comparison

Figure 4 shows the average success rates of GLT, LT and MSVM. The experimental results lead to the same conclusions as in Section 6. The advantage of GLT is obvious and as more observations are used for each time of testing, the success rate approaches 97%. We note that with GLT, it is possible to distinguish between very closely related postures which broadens the applicability of posture detection.

8. CONCLUSIONS

We considered the problem of human body posture detection with body area networks. This problem can be generalized to formation detection using RSSI measurements between wireless devices, which has other applications (e.g. formation detection of robot swarms). The problem is formulated as pattern classification problem. We developed a probabilistic (hypothesis testing based) approach, the core of which includes the construction of a pdf family representation of formation features. We further analyzed and compared this algorithm (GLT) with LT and MSVM. The simulation and experimental results show that GLT works better due to its ability to handle multiple observations and its robustness to measurement uncertainty.

In the future, more work could be conducted on the robustness of GLT and LT to get a more clear understanding of these methods in multi-observation classifications.

9. ACKNOWLEDGMENTS

The authors would like to thank Saikat Ray for useful discussions.

10. REFERENCES

[1] A. W. Bowman and A. Azzalini. *Applied Smoothing Techniques for Data Analysis*. Oxford University Press, New York, 1997.

[2] C. Burges. A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2(2):121–167, 1998.

[3] F. H. Bursal. On interpolating between probability distributions. *Applied Mathematics and Computation*, 77:213–244, 1996.

[4] C. Cortest and V. Vapnik. Support-vector networks. *Machine learning*, 20(3):273–297, 1995.

[5] K.-B. Duan and S. S. Keerthi. Which is the best multiclass SVM method? an empirical study. *Multiple Classifier Systems*, pages 278–285, 2005.

[6] D. Erdogmus, R. Jenssen, Y. Rao, and J. Principe. Multivariate density estimation with optimal marginal Parzen density estimation and Gaussianization. In *2004 IEEE Workshop on Machine Learning for Signal Processing*, pages 73–82, 2004.

[7] E. Farella, A. Pieracci, L. Benini, and A. Acquaviva. A wireless body area sensor network for posture detection. In *Proceedings of the 11th IEEE Symposium on Computers and Communications*, pages 454–459. Washington, DC, USA, 2006.

[8] F. Foerster, M. Smeja, and J. Fahrenberg. Detection of posture and motion by accelerometry: a validation study in ambulatory monitoring. *Computers in Human Behavior*, 15(5):571–583, 1999.

[9] E. Jovanov, A. Milenkovic, C. Otto, and P. de Groen. A wireless body network of intelligent motion sensors for computer assisted physical rehabilitation. *Journal of Neuroengineering and Rehabilitation*, 2(6), March 2005.

[10] C. Lai, Y. Huang, H. Chao, and J. Park. Adaptive body posture analysis using collaborative multi-sensors for elderly falling detection. *IEEE Intelligent Systems*, January 2010.

[11] B. Latre, B. Braem, I. Moerman, C. Blondia, E. Reusens, W. Joseph, and P. Demeester. A low-delay protocol for multihop wireless body area networks. In *Fourth Annual International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services*. Philadelphia, Pennsylvania, August 2007.

[12] C. Otto, A. Milenković, C. Sanders, and E. Jovanov. System architecture of a wireless body area sensor network for ubiquitous health monitoring. *Journal of Mobile Multimedia*, 1(4), 2006.

[13] I. C. Paschalidis and D. Guo. Robust and distributed stochastic localization in sensor networks: Theory and experimental results. *ACM Trans. Sensor Networks*, 5(4), November 2009.

[14] I. C. Paschalidis, K. Li, and D. Guo. Landmark-based position and movement detection of wireless sensor network devices. In *Proceedings of 46th Annual Allerton Conference on Communication, Control and Computing*. Monticello, Illinois, September 2008.

[15] M. Quwaider and S. Biswas. Body posture identification using hidden Markov model with a wearable sensor network. In *Proceedings of the ICST 3rd International Conference on Body Area Networks*. Brussels, Belgium, 2008.

[16] S. Ray, W. Lai, and I. C. Paschalidis. Statistical location detection with sensor networks. *Joint special issue IEEE/ACM Trans. Networking and IEEE Trans. Information theory*, 52(6):2670–2683, 2006.

[17] T. Torfs, V. Leonov, C. V. Hoof, and B. Gyselinckx. Body-heat powered autonomous pulse oximeter. In *Proceedings of the 5th IEEE Conference on Sensors*, pages 427–430, 2007.

[18] O. Zeitouni, J. Ziv, and N. Merhav. When is the generalized likelihood ratio test optimal? *IEEE Trans. Inform. Theory*, 38(5):1597–1602, 1992.

APPENDIX

A. PARZEN WINDOW PDF ESTIMATION

We applied Parzen windows to estimate the marginal pdfs of our observation data. For a set of scalar samples $\{x_1, \ldots, x_N\}$ the Parzen windows estimate for the marginal pdf is

$$\hat{f}(x) = \frac{1}{N}\sum_{j=1}^{N} K_\sigma(x - x_j), \tag{12}$$

where the kernel function $K_\sigma(\cdot)$ is a Gaussian pdf with zero-mean and variance σ^2. The parameter σ controls the width of the kernel and is known as the kernel size. We use the default σ value that is optimal for estimating normal densities[1].

B. DERIVATION OF THE GAUSSIANIZATION FORMULA (1)

Suppose that the i-th marginal of $\mathbf{y}$ is $f_i(y_i)$, with a corresponding cumulative distribution function (cdf) $F_i(y_i)$. Let $\phi(\cdot)$ denote the cdf of a zero-mean unit-variance single dimensional Gaussian variable:

$$\phi(\xi) = \int_{-\infty}^{\xi} \frac{1}{\sqrt{2\pi}} e^{-\alpha^2/2} d\alpha. \tag{13}$$

According to the fundamental theorem of probability, $x_i = \phi^{-1}(F_i(y_i))$ is a zero-mean and unit-variance Gaussian random variable. Therefore, the element-wise Gaussianization functions are defined as $h_i(y_i) = \phi^{-1}(F_i(y_i))$. The key assumption of this method is that $\mathbf{z} = (z_1, \cdots, z_d)$, where $z_i = h_i(y_i)$, is a jointly Gaussian vector with zero mean. Let its covariance be

$$\mathbf{\Sigma_z} = \mathbf{E}[\mathbf{z}\mathbf{z}^T]. \tag{14}$$

Hence, if the marginal pdfs of $\mathbf{y}$ and the covariance $\mathbf{\Sigma_z}$ are known (estimated from samples), the joint pdf of $\mathbf{y}$ can be obtained as

$$p(\mathbf{y}) = \frac{\mathbf{g_{\Sigma_z}}(\mathbf{h}(\mathbf{y}))}{|\nabla \mathbf{h^{-1}}(\mathbf{h}(\mathbf{y}))|} = \mathbf{g_{\Sigma_z}}(\mathbf{h}(\mathbf{y})) \prod_{i=1}^{d} \frac{p_i(y_i)}{g_1(h_i(y_i))}, \tag{15}$$

where $\mathbf{g_{\Sigma_z}}$ denotes a zero-mean multivariate Gaussian density function with covariance $\mathbf{\Sigma_z}$ and g_1 denotes a zero-mean univariate Gaussian distribution with unit variance.