An Activity Recognition System for Ambient Assisted Living Environments

Jin-Hyuk Hong, Julian Ramos, Choonsung Shin, and Anind K. Dey

Human-Computer Interaction Institute
Carnegie Mellon University
5000 Forbes Ave., Pittsburgh, PA 15213
{hjinh7,choonsung2}@gmail, {julian,anind}@cs.cmu.edu

Abstract. Ambient assisted living facilities provide assistance and care for the elderly, where it is useful to infer their daily activity for ensuring their safety and successful aging. In this work, we present an activity recognition system that classifies a set of common daily activities, where it is designed to be comfortable and non-intrusive, and is comprised of commercial, robust and well known devices. A hybrid model of Bayesian networks and support vector machines for activity recognition with calibration is proposed to provide a high recognition accuracy and fast adaptation for new users. On the data collected from 15 participants, we have compared our approach to other two ways of building activity recognition systems, and shown its superiority.

Keywords: activity recognition, ambient assisted living facility.

1 Introduction

Activity recognition for ambient assisted living facilities is necessary for supporting the well-being of occupants at short and long term, since there are many situations that could endanger or halt successful aging of the elderly, like falls. While this could become dangerous by itself, depending on the impact and the health state of the elder, it is worst when the person cannot stand up again by herself. Therefore, supervision or monitoring of an elderly person at risk for such problems is required to be able to recognize and provide assistance in such situations. Other less common daily activities could easily aggravate the health of the elder; for example, if an Alzheimer's patient simply forgets to take a pill [1], his condition could be exacerbated in the long term. These kinds of activities are much more subtle and harder to recognize by means of commonly used sensors like accelerometers. Another kind of situation that arises outside of the scope of commonly used sensor systems for activity recognition is exercising. While for most people, exercising seems standard and harmless, it is known in the medical sports community that increasing one's maximal heart rate by more than 75% increases the chance of a cardiovascular or pulmonary attack [2], and that this likelihood increases with age, particularly when other cardiovascular problems like high blood pressure are present.

S. Chessa and S. Knauth (Eds.): EvAAL 2012, CCIS 362, pp. 148–158, 2013.
© Springer-Verlag Berlin Heidelberg 2013

In this paper we present a system for real-time recognition of daily activities by using a wearable sensing platform including accelerometers worn by users. The activities that will be recognized in our system are comprised of: Walking, standing up, sitting down, lying down, bending, falling and bicycling (using a stationary bike). A hybrid model is used for precise activity recognition, and a calibration-based model construction approach is proposed for the easy setup of the system.

2 Activity Classification Problem

Activity classification is well known and it has been studied thoroughly in the area of ubiquitous and mobile computing. Many systems have been created by using accelerometers located at different parts of the body [3–5]. However, few of them use a combination of physiological signal sensors, accelerometers and localization devices (GPS, radio tomography). The setup with the three kinds of sensors can improve the classification of activities that have a higher physical demand, since physiological signals such as heart rate, temperature and respiration change according to the physical activity [6], still physiological sensors alone tell only part of the story. For example, an increase of heart rate does not necessarily imply engagement of the subject in physical activity. Instead, it could be pure emotion triggered by watching a movie [7]. In the other hand the user may be engaged in a demanding Yoga session where accelerometer's readings will not easily tell apart that there is a physical activity going on. To tackle this problem, a localization system using radio tomographic imaging (RTI) [8–10] has been included, this system can tell with an accuracy of +/- 20 cms the location of one person or more in its area of coverage. With this localization system, and a map of the living environment of the user, it will help the classifiers to discern the context of the sensor data captured by the Bioharness.

Furthermore, conventional recognition systems sometimes use an excessive number of features, but some of them are unnecessary or redundant [11] in the sense that they do not necessarily help in the classification task but instead lead to increased computation time, implying a slower system response, increased memory usage and a shorter battery life for a mobile device performing real-time analysis and sensing. In this work, we propose a system centered on overcoming all of these problems by evaluating features and selecting those relevant to the problem and to every activity. A hybrid model composed of Bayesian networks and support vector machines is exploited to model the relationship between sensory signals and activity. In modeling activity, we consider the issues of accuracy, responsiveness and comfort or wearability of the system by keeping at a minimum the quantity and complexity of data being processed.

3 Hardware/Sensing Platform

Our system is comprised of one popular physiological signals recording device, a novel system for indoor/outdoor localization and a laptop/smartphone for data logging and processing. The signals used are: electro cardiograph, breathing (chest expansion and compression), skin conductivity, skin temperature and relative localization coordinates(x,y). From these, different features are extracted (Table 1). In previous studies [6-7], these features have proven to help in the classification problem we are trying to solve.

Table 1. Features used in our activity recognition system

Signals	Features extracted	Sampling rate
Electrocardiogram	Heart rate, ECG amplitude, ECG noise, statistics (Std, mean, max, min)	250 Hz
Respiration	Breathing rate, breathing wave amplitude, statistics	18Hz
Temperature	Skin temperature, statistics	
Acceleration	XYZ acceleration minimums and peaks, posture, vector magnitude, peak acceleration, statistics	1Hz 100Hz
Coordinates	X, Y position in the living area	10Hz
Noise	Measurement of the reliability of the system generated coordinates	10Hz
Acceleration	XYZ acceleration statistics	20Hz

As shown in Fig. 1, our system consists of a Bioharness BT, a RTI system and a laptop/smartphone. The laptop/smartphone is used only to process the data from the sensing systems, extract features and perform activity classification. The Bioharness BT is a chest wearable elastic strap capable of measuring several physiological signals as well as physical quantities: heart rate, breathing rate, acceleration and skin temperature. Furthermore, it is also capable of wireless transmission of the logged data through Bluetooth. This device has been used before by our team to recognize stressful [12] situations and physical activities. It has outstanding electronic properties like long lasting battery life of up to 18 hours. The sensing frequency varies among the different physiological signals but provides at least a sample per second and at most 250 samples per second for the ECG.

Radio tomographic imaging [8–10], is a novel technique that measures the variation of the signal strength of a network of radio devices for localization purposes. RTI works by placing a network of small and inexpensive radios around an area of interest. Each radio communicates with the others in the network creating a dense net of links passing through the area. Objects moving within the area will either reflect or absorb the wireless signal creating a measurable disruption that is used to measure the location of the interfering object. A laptop/smartphone will be used as the main data processing device. On it, an application will be running at all times, logging the data sent by the Bioharness and the Radio tomographic imaging system. Data logging, visualization and classification will be performed on this main application. The laptop and the smartphone will be running on Windows and Android OS, respectively.

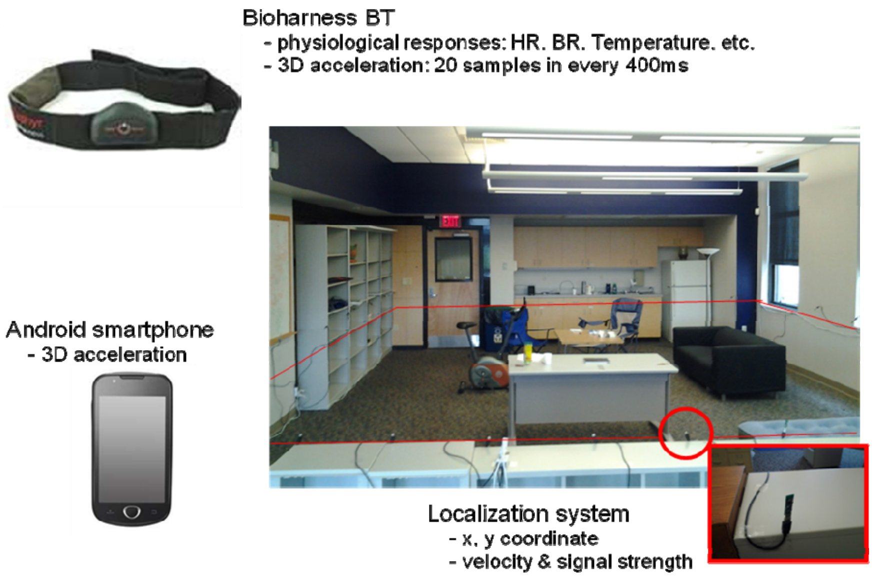

Fig. 1. Sensing platform

4 Activity Recognition System

As shown in Fig. 2, the system is composed of three phases including modeling, calibration, and online recognition. For practical use of the system, we propose a new approach to personalize the system instead of training an activity model for new users. In the modeling phase, we build a pool of activity models from a group of people. 3D acceleration data from the bioharness are collected through the android smartphone where videos are captured for the manual labeling of the data. An activity model composed of Bayesian networks and support vector machines is constructed through the process of preprocessing and model training. The preprocessing module filters out noise, segments the data every half second, and extracts statistical features (minimum, maximum, average, median, standard deviation) from the 3D acceleration.

Instead of conventional approaches that construct a population model or train an individual model for a new user, we apply a most matching activity model in the pool of activity models already constructed from the other people. The new user is only required to perform each activity shortly for the calibration process by using a simple labeling interface on the Android smartphone. In the calibration phase, it calculates the recognition performance of activity models on this new data, and selects one with the highest accuracy. Finally, the system recognizes the user activity in 2 stages by integrating two additional modules of the localization system and the accelerometer of the Android. The hybrid model of Bayesian networks and support vector machines with the 3D acceleration of bioharness, which is selected from the calibration process, first carries out the activity recognition. This recognition result is refined by several rules that are designed with the localization system and the additional 3D acceleration of the Android across the population.

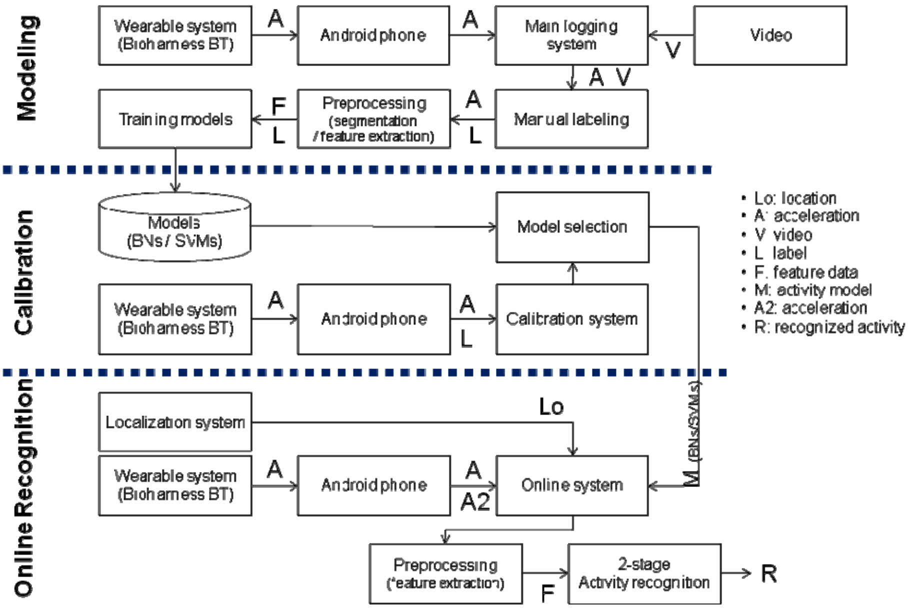

Fig. 2. Overview of the general system

4.1 Hybrid Model for Activity Recognition

As shown in Fig. 3, our hybrid model to provide a preliminary recognition result has a number of expert models, each of which is optimized to recognize the corresponding activity, and a control model that regulates those multiple expert models to output a final decision. For the given segments of extracted features, the control model estimates the probability of all activities and orders the expert models according to the probabilities. As its subsumption architecture [13], the expert models perform the classification of activities in order until an activity is recognized. This modular architecture leads to accurate modeling of each activity with an expert model, as well as effective management of those multiple expert models by the control model. Also, the proposed method has been verified its usefulness to address the ambiguity of integration of multiple models through several previous works on multiclass problems [14,15].

In constructing this hybrid model, we use Bayesian networks as the control model and support vector machines (polynomial kernel with degree = 3) for expert models to find the complex relationship between the features and activities, accompanied with a feature selection process since not all features are useful for recognizing each activity.

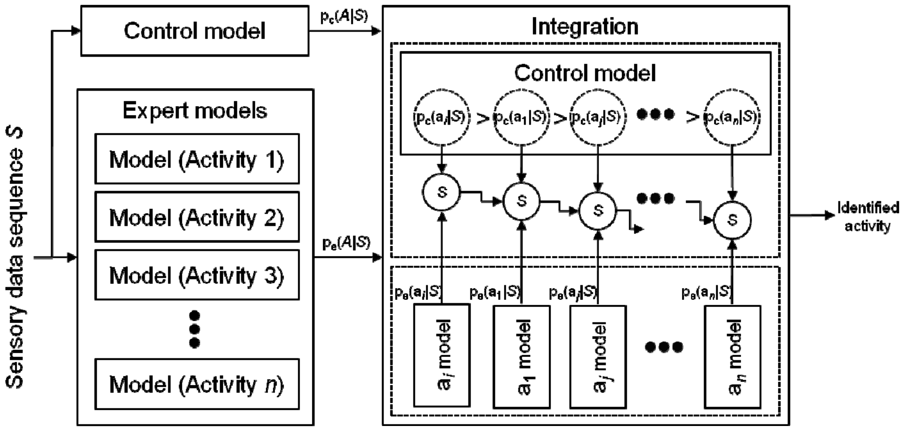

Fig. 3. Classification scheme

4.2 Additional Refining Process

After getting the preliminary recognition result from the hybrid model, we apply several rules based on the information from the Android smartphone and localization system. The rules are designed based on the general characteristics of activities and the additional sensory information as shown in Fig 4.

Localization information-based

IF UserLocation isn't around BIKE, it is impossible to bicycle
IF UserLocation is around CHAIR or COUCH, it is more probable to sit down
IF UserLocation is around BED or COUCH, it is more probable to lie down
IF UserVelocity $< \theta 1$, it is less probable to walk

Android smartphone acceleration-based

IF avg(ZX) $> \theta 2$, it is less probable to stand up or walk
IF avg(ZX) $< \theta 2$, it is less probable to sit down or lie down
IF std(XYZ) $> \theta 3$, it is more probable to walk or bicycle

For two different types of bending activities

IF Type 1 bend: avg(ZX) $< \theta 2$ &
IF 1st stage result$\in${bend, lie_down},
THEN it could be bending
IF Type 2 bend: avg(ZX) $> \theta 2$ &
IF 1st stage result$\in${stand, sit_down},
THEN it could be bending

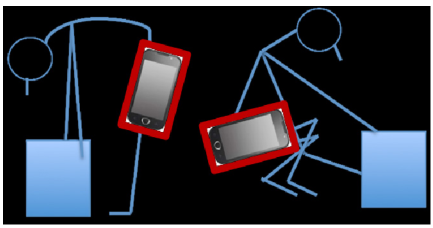

Fig. 4. Refining rules

The localization system provides the user's indoor location and its related contextual information on the indoor environment. Some specific activities are limited for the user to perform, e.g., bicycling is only available around the stationary bike and lying down has more chances when the user is on the bed or couch. The walking activity surely has certain speed due to the user's movement. The additional acceleration information of the Android in the user's pocket may help an ambiguity between some activities such as standing up and sitting down, which are often hard to differentiate based on the chest movement only. The recognition can be improved by considering the lower and upper body movement together.

5 Experiment and Analysis

5.1 Data Collection

For the validation of our system, we have collected the data in a furniture room by using our sensing platform, especially with the Bioharness. Data collection has been carried out in a large room with various furniture including a desk, different types of chairs, a couch, bookshelves, a bed, a table, a stationary bike, a sink and a refrigerator, where there are also various objects such as books, golf and tennis balls to provide a more realistic data collection environment. A camera was installed in the room to capture the video of participants' performing activities to get the ground truth through post labeling.

During data collection, participants were asked to wear the Bioharness and perform a set of tasks for about an hour. For more realistic movements, any explicit instruction or restriction was not given to the participants so that they could freely complete the tasks. 30 different tasks given to the participants include common daily activities such as walking around the room, sitting on the chairs and reading a book, lying down on the couch and taking a rest, rearranging books in the bookshelves, picking up golf or tennis balls, bicycling on the stationary bike, falling down onto the bed, having a small meal in the couch and washing hands. Especially for a falling task, a short demonstration was given by the experimenter and the participants were asked to fall down onto the bed for safety. These tasks were specially intended as a way to make the participants execute them in a naturalistic way in which they were not restricted by time or form constraints, neither they were aware of the tasks we intended to gather. After collecting the sensory data, we analyzed the video captured during the data collection and obtained the ground truth data in terms of 7 activities including standing up, sitting down, lying down, walking, bending, bicycling, and falling.

Fifteen people were recruited for the data collection, where total 112,592 samples were collected. As shown in Fig. 5, three common activities including standing, walking and sitting account for more than 20 percent, but some activities like bending and falling account for very little since they are very short activities.

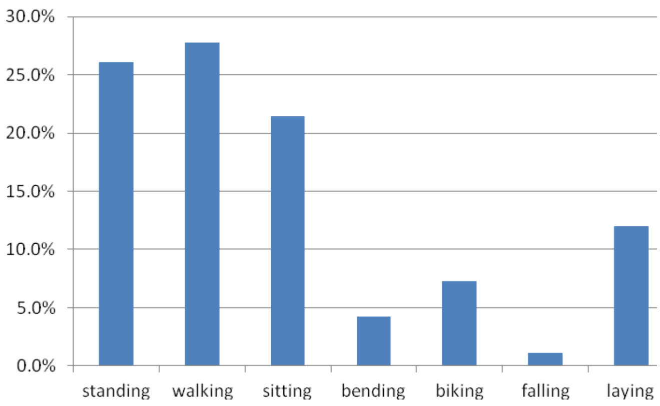

Fig. 5. Activity class distribution

5.2 Experimental Result

In 2-fold cross validation, we have compared our system with two other approaches of building activity models: population and individual. On each run for a participant, the population model was trained with the data from other people (14 participants in this study) and one fold data for training from the participant, where the individual model was built with his/her one fold data for training only. Our calibration-based approach first constructed 14 individual models from 14 other people, measured the prediction accuracy of these models on the one fold data for training of the participant, and then selected a model obtaining a highest accuracy.

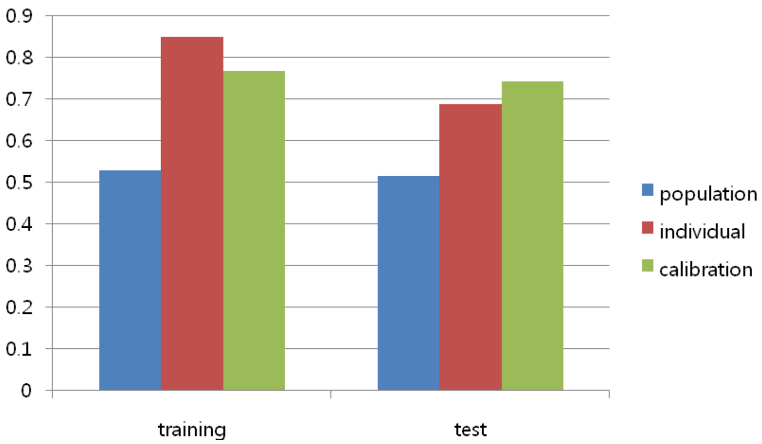

Fig. 6. Recognition accuracy

Fig. 6 shows the recognition accuracy of the three approaches for 15 participants. Due to individual variation, population models failed to achieve high accuracy not only for the test data but also for the training data. Individual models could fit to the training data by achieving an accuracy of 85% where they only obtained an accuracy of 69% on the test data because of the lack of training data. However, our calibration-based approach could get the highest accuracy of 74% on the test data among the three approaches without any additional training data. This result signifies that our approach can be applied to new users without collecting much training data from themselves to build new activity models but only relying on a small amount of data for calibration.

Fig. 7 shows a comparison of three approaches in terms of individual performance. For most participants, our approach could find a good activity model from the pool of activity models learned from other people, but sometimes it was worse than individual models since it failed to find a suitable one out of 14 other activity models. In order to get a better performance with our approach, it is required to train various models from a larger population. On this matter, we need a further investigation on each individual to understand the variation of activities across the population.

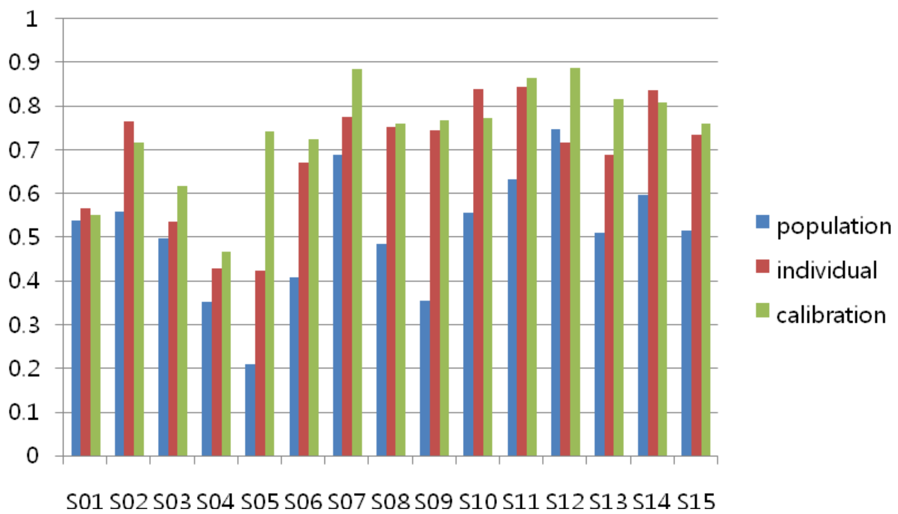

Fig. 7. Individual performance

5.3 Discussion

In our experimental result, we presented the comparison of three approaches in terms of model construction where our calibration-based approach was the best in accuracy. As shown in Table 2, however, some activities were hard to classify due to their similarity in terms of chest movement, such as standing vs. walking (8.7% errors in total), and standing vs. sitting (3.4% errors in total). Especially, many activities were confused by walking (8.9% errors in total), and bending was also misclassified by other activities many times (45.8% errors in bending). Many of the errors were occurred due to the limitation of one modality of chest movement, which could be addressed by applying the refining modules based on two types of information from the localization system and Android smartphone (see the section 4.2). For example, standing and sitting could be discriminated from each other by considering the 3D acceleration information of the smartphone. The walking activity should show a speed by the localization system, otherwise a user activity could be hard to be regarded as walking.

Table 2. Confusion matrix (st: standing, wa: walking, si: sitting, be: bending, bi: bicycling, fa: falling, ly: lying).

		Predicted						
		st	wa	si	be	bi	fa	ly
Targeted	st	73.7%	18.1%	2.9%	2.8%	1.1%	1.0%	0.4%
	wa	14.4%	77.2%	2.9%	1.5%	0.8%	1.5%	1.7%
	si	12.2%	8.8%	70.1%	1.2%	1.6%	0.7%	5.3%
	be	9.0%	18.3%	2.6%	54.2%	6.1%	5.2%	4.7%
	bi	7.7%	12.6%	1.0%	6.7%	69.6%	1.9%	0.5%
	fa	4.1%	23.1%	3.6%	6.5%	3.5%	51.4%	8.0%
	ly	1.1%	3.1%	0.6%	1.2%	1.3%	1.3%	91.5%

Our activity recognition system has been submitted to the EVAAL competition 2012 (http://evaal.aaloa.org/), and it produced an accuracy of 72% for a unknown actor who performed 7 activities in the The CIAmI Living Lab (http://www.ciami.es/valencia/), which is the highest accuracy obtained among the participating teams. On the competition, our system was only performed with the Bioharness and Android smartphone, and the localization system was excluded from the evaluation because of its malfunction by some wireless interference from the environment.

6 Conclusion

In recent, ambient assisted living facilities provide assistance and care for the elderly, where activity recognition is one of key components of the facilities. In order to address several challenging issues in activity recognition, in this paper, we presented an activity recognition system working with multiple sensors to correctly recognize 7 types of activities, where we obtained an accuracy of 74% with the hybrid model of our activity recognition system. For the easy setup of the system for new users, we proposed a calibration approach that only requires a small amount of calibration data. Besides the bioharness, we also included a localization system and an Android smartphone as other modality to address the limitation of modeling activities based on chest movement. Through an experiment and the participation of the EVAAL competition, we have shown the superiority of the proposed activity recognition system.

Acknowledgement. This material is based upon work supported by the National Science Foundation under Grant Numbers: CNS-0910878 and CNS-1035152, funded under the American Recovery and Reinstatement Act of 2009 (Public Law 111-5).

References

[1] Matthew, L., Dey, A.K.: Reflecting on pills and phone use: supporting awareness of functional abilities for older adults. In: Proc. of the 2011 Annual Conf. on Human Factors in Computing Systems, pp. 2095–2104 (2011)
[2] ACSM 2009 Guidelines for exercise testing and prescription
[3] Bao, L., Intille, S.S.: Activity Recognition from User-Annotated Acceleration Data. In: Ferscha, A., Mattern, F. (eds.) PERVASIVE 2004. LNCS, vol. 3001, pp. 1–17. Springer, Heidelberg (2004)
[4] Maurer, U., Smailagic, A., Siewiorek, D.P., Deisher, M.: Activity recognition and monitoring using multiple sensors on different body positions. In: Int. Workshop on Wearable and Implantable Body Sensor Networks, pp. 113–116 (2006)
[5] Ravi, N., Dandekar, N., Mysore, P.: Activity recognition from accelerometer data. In: Proc. of the National, pp. 1541–1546 (2005)
[6] Li, M., Rozgic, V., Thatte, G., Lee, S.: Multimodal physical activity recognition by fusing temporal and cepstral information. IEEE Transaction on Neural Systems and Rehabilitation Engineering 18(4), 369–380 (2010)
[7] Lisetti, C.L.: Using Noninvasive wearable computers to recognize human emotions from physiological signals. EURASIP Journal on Applied Signal Processing, 1672–1687 (2004)
[8] Wilson, J., Patwari, N.: Radio tomographic imaging with wireless networks. IEEE Transactions on Mobile Computing 9(5), 621–632 (2010)
[9] Zhao, Y., Patwari, N.: Noise reduction for variance-based device-free localization and tracking. In: 2011 8th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks, pp. 179–187 (2011)
[10] Wilson, J., Patwari, N.: See-through walls: motion tracking using variance-based radio tomography networks. IEEE Transactions on Mobile Computing 10(5), 612–621 (2011)
[11] Guyon, I.: An Introduction to Variable and Feature Selection. Journal of Machine Learning Research 3, 1157–1182 (2003)
[12] Hong, J.-H., Ramos, J., Dey, A.: Understanding physiological responses to stressors during physical activity. In: Proc. of the 14th Int. Conf. on Ubiquitous Computing (2012)
[13] Brooks, R.: A robust layered control system for a mobile robot. IEEE Journal of Robotics and Automation I, 14–23 (1986)
[14] Hong, J.-H., Min, J.-K., Cho, U.-K., Cho, S.-B.: Fingerprint classification using one-vs-all support vector machines dynamically ordered with naive Bayes classifiers. Pattern Recognition 41(2), 662–671 (2008)
[15] Hong, J.-H., Cho, S.-B.: A probabilistic multi-class strategy of one-versus-rest support vector machines for cancer classification. Neurocomputing 71(16-18), 3275–3281 (2008)

All in-text references underlined in blue are linked to publications on ResearchGate, letting you access and rea