Feature Selection and Activity Recognition from Wearable Sensors

Susanna Pirttikangas[1,2], Kaori Fujinami[2], and Tatsuo Nakajima[2]

[1] Department of Electrical and Information Engineering, University of Oulu, Finland
[2] Department of Computer Science, Waseda University, Japan
msp@ee.oulu.fi, {fujinami, tatsuo}@dcl.info.waseda.ac.jp

Abstract. We describe our data collection and results on activity recognition with wearable, coin-sized sensor devices. The devices were attached to four different parts of the body: right thigh and wrist, left wrist and to a necklace on 13 different testees. In this experiment, data was from 17 daily life examples from male and female subjects. Features were calculated from triaxial accelerometer and heart rate data within different sized time windows. The best features were selected with forward-backward sequential search algorithm. Interestingly, acceleration mean values from the necklace were selected as important features. Two classifiers (multilayer perceptrons and kNN classifiers) were tested for activity recognition, and the best result (90.61 % aggregate recognition rate for 4-fold cross validation) was achieved with a kNN classifier.

1 Introduction

Our system of identifying the user of a sentient artefact was presented in [3]. The system recognised the user's activity from wearable sensors (on the right wrist and thigh, triaxial acceleration signals) with a neural network and combined the information of the state-of-use of an artefact to an activity with a linkage condition (sitting activity linked to a chair, walking activity linked to a door), and in this way identified the wearer as the user of the artefact. In the case of multiple users doing the same activity, the identity of each user was verified with correlation analysis between the artefact's and user's acceleration signals. In the experiment, we studied the activities typing, sitting and walking from two different users.

This paper expands our research and explores the activity recognition of 13 different users and 17 different activities from wearable sensors as a classification problem. We describe our sensor device and some data collection issues. The goal is to find the best discriminative features that will be fast to calculate to enable prompt processing of information from multiple sources. Our sensory device can be programmed so that only the essential data is sent to a data analysing terminal. This setting requires fast processing, since the amount of data and traffic is very high as most artefacts and users will be equipped with sensor devices.

H.Y. Youn, M. Kim, and H. Morikawa (Eds.): UCS 2006, LNCS 4239, pp. 516–527, 2006.

We calculate several features from triaxial acceleration data from four wearable sensory devices. Also heart rate mean values are calculated. The feature set is pruned with a forward-backward search [4] where we utilise a kNN classifier to select the best features. We will also test different lengths of time window sizes for feature calculation in order to find a good accuracy in recognition and minimum delay of the final answer for the user. The analysis involves testing two different classifiers, a multilayer perceptron (MLP) and a kNN classifier for activity recognition.

2 Related Work

Activity recognition from wearable acceleration sensors is a current research problem. The most comprehensive study to date is presented in [1], where 5 biaxial accelerometers were attached to the hip, right wrist, left elbow, left thigh and right ankle of a testee. The authors studied 20 different activities from 20 different users, achieving a 84 % overall recognition rate with a decision tree. For some activities, such as walking, running, and standing still, a user-specific training data is reported not to have been required to reach over 80 % recognition accuracy. Furthermore, the authors emphasise the importance of natural, non-laboratory data collection. Their approach has guided us in the data collection phase as described in section 3.2. A summary of activity recognition research can be found in their paper.

The best placements for wearable acceleration sensors was sought for in [7]. For lower body, the sensors attached to the ankle or hip (also the thigh gave good results) are the best in complex situations, such as walking up or down the stairs. The analysis on upper body sensor placement is confused by the fact that the user had to carry an iPAQ to annotate their data. In [1], the dominant wrist and thigh of the user were found to be the best placements for acceleration sensors. Based on these findings, we attached the acceleration sensors on the right and left wrists and on the right thigh. Also, we attached one sensor device into necklaces worn by the testees, as it is a natural placement for, for example, sensor embedded jewellery or name tags.

Cluster precision for evaluating the best features for discriminating activitities with acceleration data has been applied in [5]. The most discriminative features are from FFT, although different components of FFT and different time window lengths are required to separate different activities. The authors also found that variance of acceleration is more accurate than the mean value. However, their data set includes only the data of two persons in the activities of walking, standing, jogging, skipping, hopping and riding a bus. The paper also lists the most frequently used features in activity analysis from acceleration signals, including the correlations between different axes of acceleration as well as the mean and standard deviation.

Depending on the system's sampling frequence, the lag of processing the FTT and saving the results into a buffer can be quite high [8] (more than ten seconds in [8], 6.7 seconds in [1]). In our system, the delay of processing must be as short as possible, since in our experience the user expects an artefact to act as a kind of a switch to an action (imagine pressing the remote control buttons). Therefore, instead of using FFT features, we calculate mean crossing values within a time window for the study. We also test the correlation between the x and y axes of each acceleration signal and the mean and standard deviation values.

3 Data

3.1 The Sensor Device

To test our activity recognition and the utilisation of sentient artefacts, we used our own sensor node named Cookie (Figure 1 (a)). Cookie is a 50 EUR cent-sized wireless sensor node that is extensible for nine types of sensors and can communicate with any device that is capable of Bluetooth v1.1 Serial Port Profile. The main board of Cookie consists of a Bluetooth module, 2-axis linear accelerometer, compass and ambient light sensor. The extensions include e.g. a heart rate sensor, force sensor and 3-axis linear accelerometer.

Fig. 1. The Cookie attached with three axis accelerometers (a) and placements of Cookie (b)

In this experiment, four Cookies were attached, one to each of the following locations: right wrist, left wrist, right thigh and necklace, as depicted in Figure 1 (b). All the triaxial accelerometer sensors from each Cookie and the heart rate sensor on the right wrist were utilised here.

Raw data was acquired on the node (acceleration sampled 64 times at 200kHz) and the averaged value was sent to a data collecting terminal at every 100 msec at 9600bps. For tests performed outside the permises of our laboratory, the data collecting terminal was worn by the user (Sony VAIO Type-U running Windows XP SP2 with USB-Bluetooth dongle).

3.2 Semi-naturalistic Data Collection

To achieve as natural a setting as possible for the data collection, a semi-naturalistic data collection schema, as suggested in [1], was applied. The semi-naturalistic data collection aims at minimising the possible disturbance of outside observers and the awareness of perfoming some specified activity (leading to unnatural behaviour).

In [1], the users were given a worksheet describing an obstacle course of activities to be performed. The users wrote down the starting and ending time on the worksheet and, therefore, no research observation was needed. We did not, however, find the approach of using a workheet very feasible, so we built two different kinds of user interfaces to help annotate the data and describe the obstacles for the testees. These interfaces support the labelling of different activities (time stamp labels), syncronising the data and organising it.

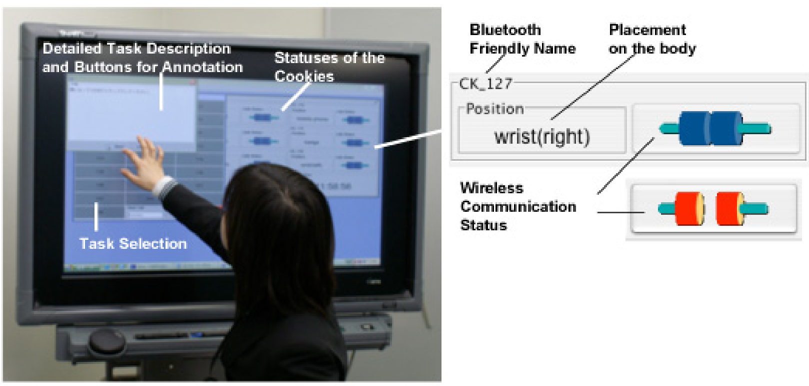

Fig. 2. The touch screen-based user interface

Figure 2 presents the obstacles in a large touch screen as tasks. When the user selects a task, a detailed description of the obstacle is given. Each task description GUI includes a button for labeling a starting time and ending time for the activity. On the right in Figure 2, the statuses of all wearable Cookies are shown. Thus, if the Bluetooth connection is lost, the testee will be able to contact the researchers and the loss of data is prevented. We also added a sound feature for the interface: a sound is played to indicate the success of time stamping each time a task is started and ended. Furthermore, after one minute, a sound is played to inform the user about the time spent in an activity. The minimum of one minute of data were collected for each activity. The user interface is bilingual (Japanese and English) and example tasks are shown in Table 1.

Table 1. Examples of the task descriptions

Task 1	Clean the white board with the brush.
Task 2	Sit down and read a newspaper. Read at least one whole article from the paper.
Task 3	Stand still for a minute
Task 4	Sit down to the sentient chair and relax
Task 5	Sit down, switch on the TV with the remote control and watch some TV
...	...
Task 14:	Walk up the stairs (at least to the 3th floor)
...	...
Task 17:	Ride around the laboratory with a bicycle

Fig. 3. The wearable user interface

A wearable user interface (Figure 3) with a slider to select the description of each task was built for tasks performed outdoors or not in the vicinity of the screen (Figure 2). The time stamps are here labelled with a toggle switch on the right side of the interface. The interface shows a short description of each task in English, and it also informs the user if the Bluetooth connection is lost.

Both interfaces record the test details in a log file, from where they are easy to extract. The placements of the Cookies, as well as the data file name associated to each Cookie, is logged. Furthermore, the time stamps for the beginning and ending time of each activity and a short description of the tasks are saved.

The data were collected from 13 testees, 9 males and 4 females, aged from 22 to 32. The users were given a brief introduction of the obstacles and the equipment, and the users performed the tasks and labelled the activities with time stamps on their own, utilising the built interfaces. In some cases, a researcher was observing the activities and in one case, a video recording was made. The testee indicated

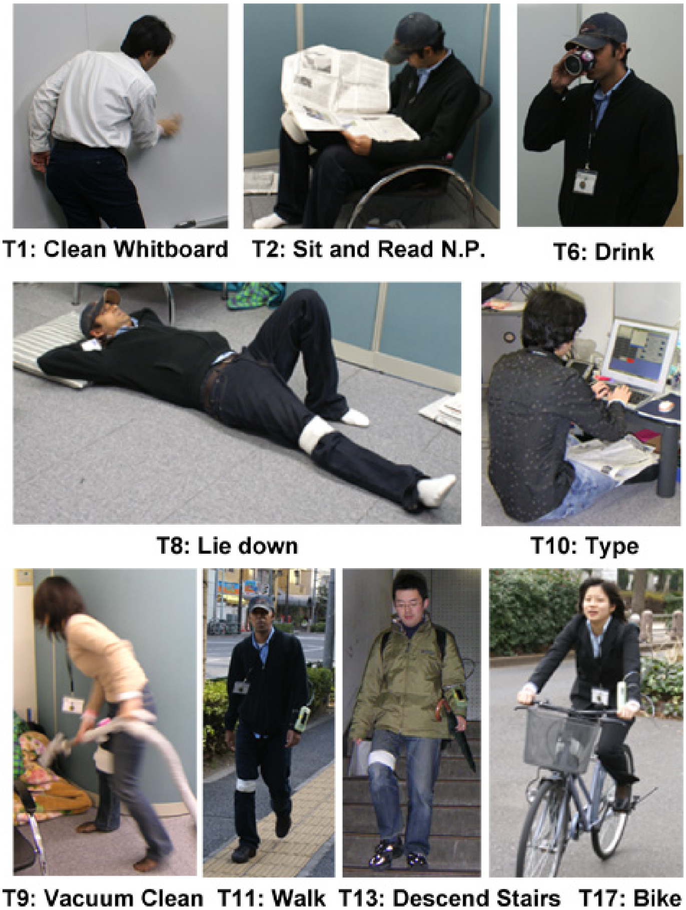

Fig. 4. Pictures from data collection tasks

that this did not affect his behaviour. Pictures from the data collection are shown in Figure 4.

3.3 Feature Selection

Regarding an online setting including several Cookies that send data continuously to a data processing terminal, the design of the activity recognition system should remain simple and the lag affected by feature calculation and traffic should be small. As described in section 3.1, the Cookie is extensible to various types of sensors and can be programmed so that not all of the data are sent. We must, therefore, select the most important information from each Cookie to be sent. The delay and traffic will decrease if we can find a small subset of the most important (in view of recognition accuracy) features.

From each Cookie, we calculated only features that are fast to process in a short time window: each axis (3D) acceleration mean and standard deviation values, correlation coefficients between the x and y axes of acceleration, and mean crossing values for each axis acceleration. Furthermore, the heart rate mean from the Cookie placed on the right wrist was calculated. The tested features are listed below:

acceleration mean (x, y, z axes)
acceleration standard deviation (x, y, z axes)
acceleration correlation between x and y axis (x, y axes)
acceleration mean crossing (x, y, z axes)
heart rate mean

The features mentioned above were tested with forward-backward search [4], which is a well-known feature selection algorithm. With this procedure, a subset of best (giving the best classification result) features can be determined for the final analysis. In forward search (FS), every feature is tested for the classification one by one, and the best is selected to a subset of best features. The features that remain are then tested with the selected one, and the best one is selected to the subset and so forth. The procedure starts from one feature. The FS finds the best single features but does not find the best combination subset.

Backward search (BS) starts with classifying all features and removing the one that is lowering the classification result. In forward-backward combination, two features are selected with FS and one is removed with BS. The classification is usually done with a simple classifier, such as kNN, which was also used in this procedure. We utilised Matlab technical language to process the data and calculate the features and the LNKnet [6] pattern recognition software for other analyses.

3.4 Testing and Training

[1] state that the individual variation in body worn acceleration may be dominated by strong commonalities between people in activity patterns. Therefore, we combined all the data from different testees into one data set. The data were divided and synchronized based on the labelled time stamps to different activities, and the evaluations of the activity recognition algorithms were done with 4-fold cross validation.

From the beginning of each activity, 50 samples (5 seconds) were removed, because in some cases the starting of an activity after reading the task description and labelling the data takes a while (e.g. moving away from the screen towards the white board). Furthermore, if the time spent on an activity exceeded five minutes, the data were truncated to get as homogenous (in view of number of examples) data set as possible. Several MLP and kNN classifiers were tested for activity recognition as described in the following. The 17 activities studied are shown in Figure 6.

4 Results

4.1 Features

Three separate feature selection tests were made. In the first experiment, the mean, the standard deviation and the mean crossing values of each acceleration signal were calculated in windows of five samples (0.5 seconds). The mean values of heart rate from the right wrist were also included in the feature set. The number of features was then 37. Forward-backward search was applied, and the final feature set consisted of 19 features. No mean crossing values nor heart rate means were in this group, but all of the mean values of each acceleration signal from each Cookie (except necklace mean acceleration y) were included.

Table 2. The selected features for activity recognition

Cookie	RightWrist	LeftWrist	Thigh	Necklace
	mean X	mean X	mean X	mean X
	mean Y	mean Y	mean Y	mean Y
	mean Z	mean Z	mean Z	mean Z
		std X		
	std Y			
	std Z		std Z	

Another experiment was made with a window size of 7 samples (0.7 seconds), and this time the correlations between the x and y axis accelerometers were included into the feature set (now 28 features). All the means from all the Cookies were present in the final subset. The standard deviations from the right wrist (y, z), left wrist y, and thigh (y, z) were also included (a total of 18 features). The correlations were not included in the subset of best features.

Finally, we made one more feature selection from the mean and standard deviation values of each Cookie's accelerometer data. This time, the window length was five and the feature set included 25 features (note that the HR mean value was tested again). The final 16 features selected are presented in Table 2. The standard deviations (y, z) of the dominant wrist (all of the testees were right handed) are included. It is expected that in the case of left handed people, the standard deviations (y, z) of the left wrist would be present, but we cannot answer this question with this data.

As the mean values of the necklace were included in all of the best subsets, we made a test (with window length of seven samples) where only the mean and standard deviation information from three Cookies, i.e. the right wrist, left wrist and thigh, were included. We compared the test results with and without the necklace features and noticed that when removing the necklace information, the classification error of "sit and read a newspaper", "sit and relax", "sit and watch TV" and "lie down" was increased considerably. The highest increase of confusion was between

"sit and read a newspaper" and "brush teeth"
"sit and read a newspaper" and "sit and relax"
"sit and read a newspaper" and "sit and watch TV"
"sit and read a newspaper" and "lie down"

This result suggests that the necklace helps in detecting certain kind of movement of the upper body while seated.

4.2 Time Window

With the selected feature set presented in Table 2, experiments were made to test a suitable time window for the feature calculation. We tested windows of 1, 2, 5, 7, 10 and 15 samples (from 100 msec to 1.5 seconds) for both MLP and kNN. The results can be seen in Figure 5. The best result was achieved with the kNN classifier for the window length of 10 (1 second), giving a 90.61 % recognition accuracy.

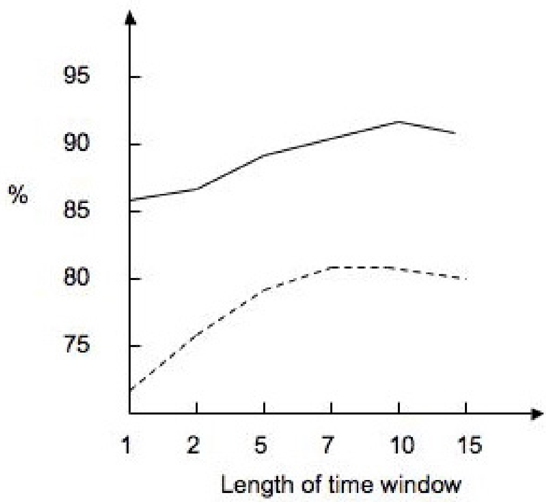

Fig. 5. Classification results in percents for the MLP and kNN for different time windows (x axis). The dashed line represents the results for the MLP.

The aggregate confusion matrix for the 17 activities and the best classification result (kNN, time window 1 second) with 4-fold cross -validation is shown in Figure 6.

4.3 Simplifying the Problem

A sentient artefact is usually developed to achieve a specific task and, therefore, the state-of-use of an artefact can be a cue to infer the user's activity [2].

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	< classified as
336	2	2	1	0	5	14	0	7	0	0	1	1	2	1	0	0	0=clean whiteboard
4	352	0	5	10	4	0	1	10	0	1	1	1	0	0	0	1	1=read a newspaper
4	0	379	0	3	1	0	0	1	0	0	0	0	1	1	0	0	2=stand still
2	7	0	370	3	3	0	2	1	1	0	0	1	0	0	0	0	3=sit and relax
1	10	1	1	357	5	0	0	12	3	0	0	0	0	0	0	0	4=sit and watch TV
2	3	1	1	2	336	4	0	8	8	3	3	0	5	1	0	1	5=drink
4	0	0	0	1	5	378	0	2	0	0	0	0	0	0	0	0	6=brush teeth
0	5	0	1	1	0	0	340	2	1	0	0	0	0	0	0	0	7=lie down
7	16	2	0	6	17	0	0	302	0	9	14	8	4	2	0	3	8=vacuum clean
0	1	0	2	0	4	0	1	0	369	0	0	0	0	0	0	2	9=type
0	1	0	0	0	1	1	0	2	0	307	30	18	0	0	0	0	10=walk
0	0	0	0	0	1	0	0	3	1	35	221	17	1	0	0	0	11=climb stairs
0	1	0	0	0	0	0	0	0	0	36	26	176	1	0	0	0	12=descend stairs
3	0	4	0	1	4	2	0	5	0	0	0	2	235	12	0	1	13=elevator Up
1	1	1	0	0	0	0	0	2	0	0	0	0	16	196	0	0	14=elevator Down
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	148	0	15=run
0	0	0	0	0	1	0	0	0	0	0	0	1	1	0	0	343	16=cycle

Fig. 6. The aggregate confusion matrix of 4-fold crossvalidation for kNN classifier based on semi-naturalistic data from 13 subjects

In our system to identify the user of an sentient artefact, the linkage condition `activity-artefact`, such as `vacuum cleaning-vacuum cleaner`, is the unifying element between an artefact and the user. However, the activity needs not to be this detailed, but could be e.g. `cleaning-vacuum cleaner`. If the linkage condition is satisfied, the more accurate activity for the wearer can then be determined from the artefact.

To simplify the recognition problem, we unified the activity classes `vacuum clean` and `clean whiteboard` to an activity `clean`. We also unified the classes `sit and read newspaper`, `sit and relax`, `sit and watch TV` and `type` into the activity `sit`. The activities `stand`, `elevator Up` and `elevator Down` are replaced simply with the activity `stand`. Climbing and descending stairs is now `stairs`. There are thus nine (9) different activities to be recognised. The scalability of the system is improved considerably if only the basic activities are modeled. This way, when new elements (artefacts) are added to the environment, only a new linkage condition needs to be added.

The best selected features and window size of seven (as no considerable improvement in classification accuracy from seven to ten samples with these features in Figure 5) samples were utilized to train a kNN and a MLP classifier to identify the nine activities. The recognition accuracies were 89.76 % and 92.89 % for the MLP and kNN, respectively.

4.4 Discussion

The features selected are the most commonly used features for activity recognition, but as a contradiction to the work in [5], the mean values of acceleration signals were the most important features instead of standard deviations. Furthermore, the correlation and mean crossing values of acceleration signals or heart rate means were not included in the subset of best features in any of the tests performed in this investigation.

The most interesting finding in feature selection was the importance of the necklace features. It is easy to embedd sensory devices to a necklace, which can also have some aesthetic value. The necklace features were found to escpecially improve the discrimination of different sitting activities (sitting and relaxing as opposed to sitting and reading a newspaper, for example). Otherwise, the importance of the thigh and the dominant wrist, as stated in [1] and [7], was confirmed.

The best recognition accuracy was achieved with a time window of ten samples (1 sec) to enable over 90 % recognition accuracy, but the results for the window sizes of five and seven were not considerably lower. When testing different time windows, the results from MLP were weaker than from kNN, but the fact that the features were selected with the latter is, naturally, an advantage to it.

In our final test, we unified the information of different activites that involved sitting into a class titled `sitting`, activities that involved standing into `standing` and activites that involved cleaning into `cleaning`. The final number of activities to be recognised was considerably lower (from 17 to nine) and the final classification result was much better (92.89 %).

5 Conclusions and Future Work

We described here the development of an activity recognition system starting from semi-naturalistic data collection to feature selection and finally proceeded to recognising different activities. The aim was to find a good set of features that enable fast processing of information giving an acceptable recognition rate.

Two different kinds of user interfaces were built for assisting the collection, synchronisation and management of data. The interfaces were found to be very useful and can be applied in any data collection task that involves a Bluetooth based device.

In our analysis, we provided a minimal feature set needed for the recognition of different activities occurring in daily life. The system's sensory device can be programmed to send only some of the data to a data analysing terminal and find the minimum set of features to enable the decrease of both computing time and traffic in the system. The features were tested with two classifiers (MLP and kNN) and the recognition accuracy for 17 activities was 90.61% in the best case. The necklace features were also found to be important in the recognition tasks. It remains to be studied whether the activity recognition algorithm for the nine basic activities studied in section 4.3 will still need these features to be present.

The ideal situation for a calm environment is that recognition accuracy could be increased up to near 100 %. To enable this, either the recognition capability of the machine learning algorithm must be made better or the other elements of the environment should be utilised in the recognition. Our vision is that the higher level context (from cleaning to vacuum cleaning, for example) will come from the artefact the wearer is using.

Our ongoing research addresses testing the combination of artefacts and wearable devices in the real world. We have already implemented the system proposed in [3]. Our future research includes the trade off analysis between the recognition accuracy and delay for the final answer. Furthermore, we are studying the situation of unknown (for the system) activity or inaccurate decision by the system and attempts to recover from such situations by rejection analysis.

Acknowledgments

This work was partially funded by the Academy of Finland, Finnish Funding Agency for Technology and Innovation and companies.

References

[1] L. Bao and S. S. Intille. Activity recognition from user-annotated acceleration data. In *Proc. Int. Conf. Pervasive Computing (Pervasive 2004)*, pages 1–17, 2004.
[2] K. Fujinami and T. Nakajima. Sentient artefact: Acquiring user's context through daily objects. In *Proc. 2nd Int. Workshop Ubiquitous Intelligence and Smart Worlds (UISW2005)*, pages 335–344, 2005.
[3] K. Fujinami, S. Pirttikangas, and T. Nakajima. Who opened the door?: Towards the implicit user identification for sentient artefacts. In *Adjunct Proc. Int. Conf. Pervasive Computing, (Pervasive 2006)*, pages 107–111, 2006.
[4] Fukunaga. *Introduction to Statistical Pattern Recognition*. Academic Press, 1990.
[5] T. Huynh and B. Schiele. Analyzing features for activity recognition. In *Proc. Joint Conf. Smart Objects and Ambient Intelligence (sOc-EUSAI '05)*, pages 159–163, 2005.
[6] L. Kukolich and R. Lippmann. Lnknet user's guide, August 1999. http://www.ll.mit.edu/IST/lnknet/ (available on 18.4.2006).
[7] B. S. N. Kern and A. Schmidt. Multi-sensor activity context detection for wearable computing. In *Ambient Intelligence*, volume 2875/2003, pages 220–232, 2003.
[8] T. Yamabe, A. Takagi, and T. Nakajima. Citron: A context information acquisition framework for personal devices. In *Proc. the 11th IEEE Int. Conf. Embedded and Real-Time Computing Systems and Applications (RTCSA'05)*, pages 489–495, 2005.