The Hearing Trousers Pocket – Activity Recognition by Alternative Sensors

Gerald Bieber[1], André Luthardt[1], Christian Peter[1,2], Bodo Urban[1]

[1] Fraunhofer IGD
Joachim Jungius Str. 11, 18059 Rostock, Germany
{gerald.bieber;andre.luthardt;cpeter;urban}@igd-r.fraunhofer.de

[2] Graz University of Technology
Inffeldgasse 16c, 8010 Graz, Austria
c.peter@cgv.tugraz.at

ABSTRACT
In daily life, mobile phones accompany the user permanently and are worn often in the front pocket of the trousers. The sensors included in today's mobile phones can hence be used for ubiquitous assistance. For instance, the acceleration sensor could be used for analysis of the person's bodily activity, or the microphone can be used to analyze the environmental noise levels. A possible sensor fusion provides additional and assured environmental and context information.

This work presents new methods of activity recognition by acceleration and sound sensors by means of sensors included in commercially available smart phones during everyday life. We could identify that sounds provide valuable additional information on a user's situation that allow to better asses a person's current context.

Categories and Subject Descriptors
H.5.2 User Interfaces (Information Interfaces and Presentation) I.5.2 Design Methodology (Pattern Recognition) J.3 Life and Medical Sciences

General Terms
Algorithms, Design, Measurement, Reliability, Human Factors

Keywords
Activity Monitoring, Acceleration, Sound Sensor, User State Detection, Context, Situation, Assistive Technologies, Physical Activity, Mental State, Stress Detection, Mobile Phone, Microphone.

1. INTRODUCTION
In modern life, sound surrounds the user in almost every situation. Even without a speech or voice recognition, sound provides additional information about the ambient situation of the user. With mobile phones being ubiquitous today, the mobile devices are carried permanently by most people, accompanying the user throughout the day, during work and in private life.

Modern phones, so called Smartphones, have a number of sensors included. Apart from the microphone and digital camera, Smartphones feature an acceleration sensor for showing pictures in the proper orientation, and GPS and sometimes even a compass for assisting their user in navigation. These sensors obviously can be used for finding out about a person's activity, context and situation. For instance, the microphone of the Smartphone can be used to record the ambient sound of a person, or the acceleration sensor can be used to detect certain levels of bodily activity.

Furthermore, today's Smartphones provide enough computing power to perform fairly sophisticated signal analysis on the device.

As has been shown in other projects [1, 2, 3], analysis of movement patterns based on acceleration data allow for automatic classification of typical everyday activities such as car driving, cycling, walking, running, jumping or other activities like house work or gardening [2]. Combining this information with other sensors' readings allows for an even better situation assessment [3].

In the project described here we investigate in the possibility of including sound information to enhance a mobile phone's activity recognition capabilities. We assume that the inclusion of sound information improve the determination of more possible activities and will enhance the accuracy of activity recognition, provided the quality of the sound information is good enough to extract relevant features even when the phone is worn in a trousers pocket.

In this paper we first describe the current work in the field of physical activity recognition by acceleration and sound sensors. In the next section we introduce a novel method of activity recognition by the combination of sound and movements. The following paragraphs describe the evaluation of the fused activity recognition and discuses the received results. The paper closes with a summary and the outlook on further research and next steps.

PETRA'11, May 25 - 27, 2011, Crete, Greece.

2. RELATED WORK

Sound analysis is well known in other disciplines such as mechanical engineering. Sound can be used to identify worn out bearings or machine parts. Bearings with imbalances produce significant distinguishable sound patterns, which can be identified by the use of microphones. Another example of using sound in engineering is for detecting lack of lubrication as this also results in different sounds. Sound classification can hence be used for scheduling maintenance work [4] and a better estimation of the current machine situation.

The effect of sound on humans in HCI settings has been explored by some researchers [5, 6], highlighting the importance of the audio channel for the affective perception of a situation. Sound affects the mood and emotional state of people as it can be recognized in shopping malls, airport or waiting zones [7]. The ambient sound probably correlates to the emotional state of the user.

Acceleration sensors in mobile phones have been used for fall detection of elderly people [8]. While this has been done successfully in the lab, it was also concluded that for real-life application the detection rate has to be improved. Using additional sensors such as the microphone has been suggested.

The combination of an accelerometer and a microphone for detecting activities in everyday life is evaluated e.g. within the MiThril project of the MIT, USA [1] and Palantir, Finland [9]. Hereby a complex sensor infrastructure (shirt, rucksack etc.) was used for the determination of sensor data. Activity recognition systems based on an accelerometer which is built into a commercial mobile phone is provided by DiaTrace [2]. Hereby the device can only distinguish between movement patterns. Further, DiaTrace cannot distinguish if a person is mentally active or not e.g. sleeping or in discussion with other people. Obviously, sound analysis is an interesting option here.

While sound analysis with body-worn or distributed sensors is widely known in ambient surroundings, we could not identify significant research about utilizing an ordinary mobile phone as sensor worn in a pocket with the purpose of using the combination of sound and acceleration sensor for physical activity recognition.

3. Physical Activity Recognition

3.1 Acceleration Data Analysis

Activity recognition from acceleration data has been done in several projects [1, 2, 9]. Their results can be summarized that physical activity recognition is possible with just one high performance acceleration sensor, at least in laboratory environments. The challenge now is the development of a suitable method of preprocessing and the identification of relevant features for activity recognition in everyday life. The DiaTrace project [2] seems to be the only project run for years, with steady improvements and evaluations in real-world settings [3, 10, 11].

The DiaTrace system used in this project allows the identification of physical activity in everyday life on a standard mobile phone. A 3D-acceleration sensor integrated in a Smartphone is used to determine physical activity by domain specific feature extraction. By use of data mining techniques and a preprocessing of acceleration data, suitable features can be extracted describing certain physical activities. The proof-of concept prototype receives a recognition rate of nearly 90 % of the activity types resting, walking, running, cycling and car driving, just by wearing the device in the front pocket of a trouser. For more detailed information on the DiaTrace project please refer to [2].

3.2 Sound Analysis

The purpose of sound analysis is to use the standard built in sound sensor (microphone) of a commercial mobile phone to improve activity recognition. While the human being is able to recognize sound in the spectrum of 20-20.000 Hz, a mobile phone is optimized to record speech data within a frequency range of ca. 100 Hz-8 kHz. The microphone of a commercial phone provide sufficient sensitivity to receive spoken voice data within a room and can even sense noise from far distances as it might affect the user.

The typical wearing position of a mobile phone is in the trousers pocket [12]. Hereby the trousers are damping and filtering the signal and weak signals will be fully absorbed. To identify most common damping situations, we produced a test sound of 900 Hz and determined the intensity of the received signal depending on different situations.

We defined the sound intensity as 100% when the sound source is in distance of 20 cm to the phone, without having any obstacles or signal weakening objects in between. We found out that the sound amplitude is 66%, when the Smartphone is in a closed bag (e.g. Rucksack, handbag), but nearly 100% when the Smartphone is in the trousers front pocket. This leads to the assumption that a trousers pocket does not significantly reduce the signal strength. However, the influence of the trousers pocket to the sound is only moderate while the user is not moving. When the user is performing physical activity, the phone is moving inside the pocket and the trousers' garment is rubbing on the microphone, which modifies the sound signal significantly. It is hence assumed that the quality of the signal will be affected insofar as that the potential of extracted features from the signal might be limited.

Figure 1 shows the acoustic signal and the frequency spectrum collected from a person while walking and working with a computer, respectively. During the activity walking (which was performed in a quiet environment) the phone received significant more intense sound signals than with the activity working with a computer (which was done sitting on a chair). The sound analysis provides features which diverts in sound amplitude, sound energy and sound spectrum. Beside quiet activities, some activities are performed under loud conditions such as car driving (noise by engine, wind etc.) or watching TV.

As can be seen, there are significant differences to be observed. We believe that there are many more activities which can be distinguished by analyzing the frequency spectrum of a sound signal. To receive a higher accuracy of activity recognition, we assume that combining information from different sensors will provide better classification results.

3.3 Fusing Acceleration and Sound Data

An a new approach, we use activity data from and accelerometer and sound information collected with a microphone to better recognize activity patterns of persons. Both sensors are built into a Smartphone. For activity recognition only by acceleration sensor, the DiaTrace activity recognition system developed by Fraunhofer IGD is used [2]. In order to improve the detection rate and to extend the detectable activities, the environmental sound around the user is analyzed using the Smartphone's microphone as sensor.

By fusing the accelerometer data with the sound data, an improved activity estimation should be possible such as detecting additional activities and to detect the intensity of activities.

As a first additional activity to be detected with the help of sound analysis, office work has been chosen. While there is not much bodily activity, sound analysis could give valuable hints on the person's current doing. For example, mouse clicks and keystrokes can easily be detected by a phone when lying on the desk (figure 1). Also it is possible to distinguish the activity-related sounds from other sounds in the frequency spectrum, such as driving a car or cycling (figure 2). The acoustic signal when driving a car is clearly different to the sound data of cycling. Generally we can concretize activities like physical rest, e.g. when the user only works with the hands as is the case with many office workers.

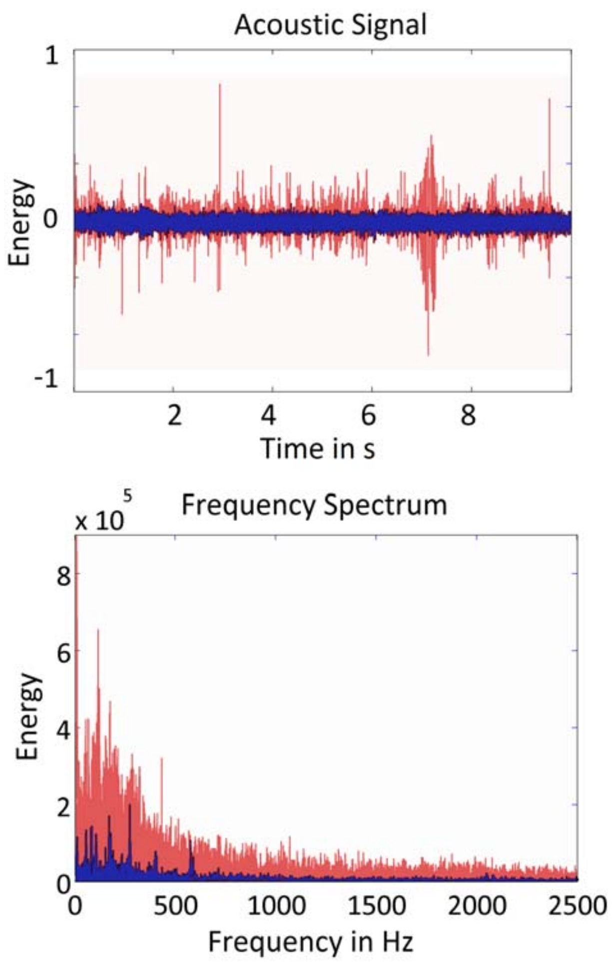

Figure 1: Noises in trousers pocket while walking (red) and working with a computer (blue). *(1) The first diagram shows the acoustic signal over time. (2) The second one the frequency spectrum of the acoustic signal.*

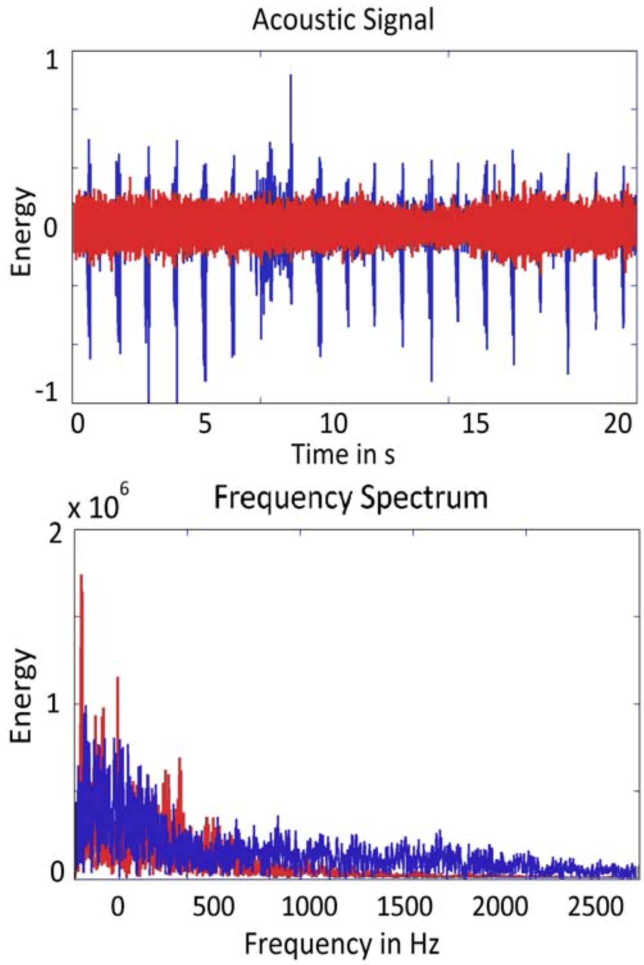

Figure 2: Noises in trousers pocket while motoring (red) and cycling (blue). *(1) The first diagram shows the acoustic signal over time. (2) The second one the frequency spectrum of the acoustic signal.*

However, it is interesting detecting activities which cannot be detected by an accelerometer. Within an office, the physical activity is mostly sitting, while the mental load might vary over the day. A discussion or a meeting with another person can be mentally demanding or even stressful, which cannot be detected with accelerometer-based activity detection systems. A sound sensor might help to detect those different situations.

For detecting relevant sound patterns, the sound data can be analyzed for periodic events. In the frequency spectrum we can clearly see features which appear periodic.

ARCHITECTURE

Based on the architecture of the accelerometer-based DiaTrace system we extend its functionality by adding a sound sensor. Both sensors provide raw data which are to be preprocessed. After a signal reconstruction, the sensor selection module decides if feasible sensor data are available.

Usually the signals of both sensors are then used for building the feature vectors for each signal, with the feature vectors describing all known relevant and extractable characteristics of the instance.

A classifier finally determines a predicted class of activity. In our work we are using a hierarchical classification algorithm. For each sensor type we are extracting a specific feature vector and classify for respective classes.

The results of the classifications are fused by a second level classifier. This classifier also uses context information for reasoning about possible activities to improve accuracy of the recognition. For instance, it can be assumed that. immediately after car driving no cycling will performed because some walking must be in between. Figure 3 shows the general architecture of the adapted DiaTrace classification kernel which is suitable to run on a mobile phone in real time.

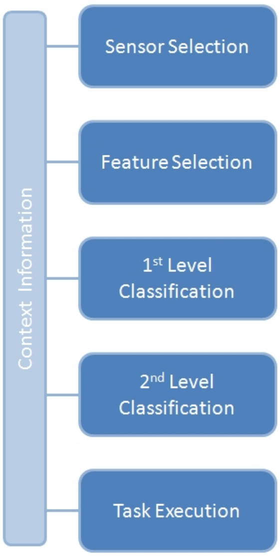

Figure 3: Generic architecture for motion tracking

4. PROTOTYPE

The presented concepts were implemented on an android smart phone, Sony Ericsson Xperia X10 with OS Android 2.2, 1 GHz. This mobile device provides an acceleration sensor, using Bosch Sensortec BMA150. The microphone is a standard electret condenser microphone. The 3D acceleration data are sampled by 32 Hz, the sample frequency for the sound data is 8kHz.

The prototype application encloses data acquisition of acceleration and sound, feature extraction (incl. Fourier analysis) and classification. Hereby we are using a decision tree as the classifier as this has been identified as a robust classifier providing good classification results while having marginal requirements on the computational power of the mobile phone. The Smartphone can be used with its normal functionality while the prototype application runs in a background modus.

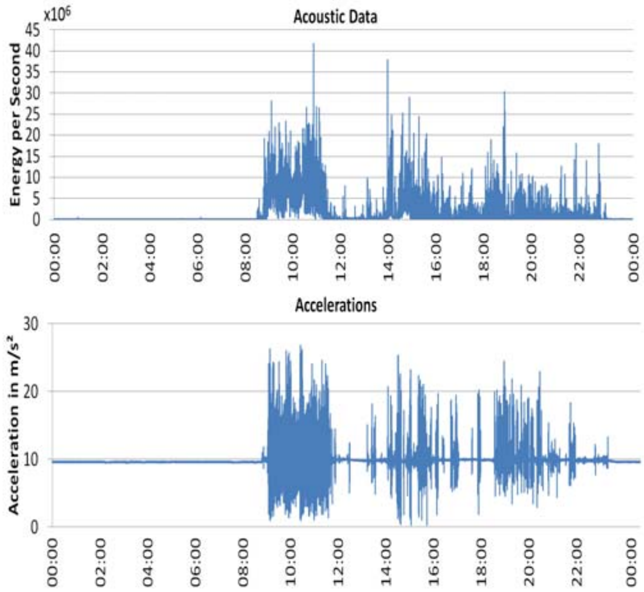

Figure 4: The sound volume and accelerations recorded of a whole day. *The subject sleeps to 9 am. Following he is cycling two hours. After that the subject relaxes to 2 pm. In the next three hours he is motoring. At the end the subject does indoor work.*

Figure 4 illustrates the high correlation between the physical activity as determined by the acceleration sensor, and the correlating sound data as determined by the microphone, over a period of 24 hours in an everyday setting. The graph shows a strong correlation between sound and acceleration.

Feature extraction is implemented separately for each type of data, including signal reconstruction and filtering. Before extracting features from the sound data, the acoustic signal is segmented into sequences with a window size of 4096 samples. Using the second power, a frequency analysis is performed using Fast Fourier-Transformation (FFT). Before transforming the acoustic data into the frequency spectrum, it has to be analyzed first whether the amplitudes are sufficiently large. This way sound was only recorded and analyzed when data of sufficient quality were available and only then features were extracted.

In the *first level classification* the sound intensity is determined. Assuming that the stress level of a person in a classical work environment increases with the intensity of environmental sound, regardless the physical activity, the sound intensity is an important factor for estimating the stress level of a person. Assuming also that increased bodily activity is accompanied by increased noise, an increased sound intensity might also be indicative for increased physical activity.

In the next steps, the activity recognition system will compute the FFT and analyzes the frequency spectrum. The FFT can identify frequencies between 1Hz and 2048Hz with its magnitudes. One relevant feature for physical activity classification is the frequency with the highest amplitude, called maximum frequency. The maximum frequency is an indicator for specific activities.

During the same time of feature extraction and classification of sound data, the mobile phone analyzes sensor data from the accelerometer. The data of the acceleration sensor consists of a 3D tuple of an orthogonal force vector. These acceleration samples are segmented analogue to the sound data. One segment includes 64 acceleration samples and is transformed in a next step into the frequency domain using FFT. Now, the system can extract features which allow classifying activities.

Interestingly, the frequency spectrums of most activities vary, but there are activities which have nearly the same frequency spectrum. Inclusion of sound data might now help to decide which activity the user performs. The results of the first level classification will be evaluated in the second level classification.

In the *second level classification* it is decided which activity is performed by the user in which intensity.

Figure 1 gives an example of the frequency spectrum and the acoustic signal of a person working with a computer. Clearly to identify is the distribution of the energy in the lower frequency band and the peaks. Every peak is symbolized by a specific frequency. A peak occurs not only once, but several times in one segment. So a peak can be a specific mouse click or a keystroke while the user is working with the computer. The shown frequency spectrum is typical for office noises.

The first prototype of sound and acceleration data analysis, implemented as a mobile app, requires an additional power consumption of about 25%. We assume that energy saving options will reduce this parameter significantly. A constant monitoring of physical activity should hence be possible with a standard Smartphone.

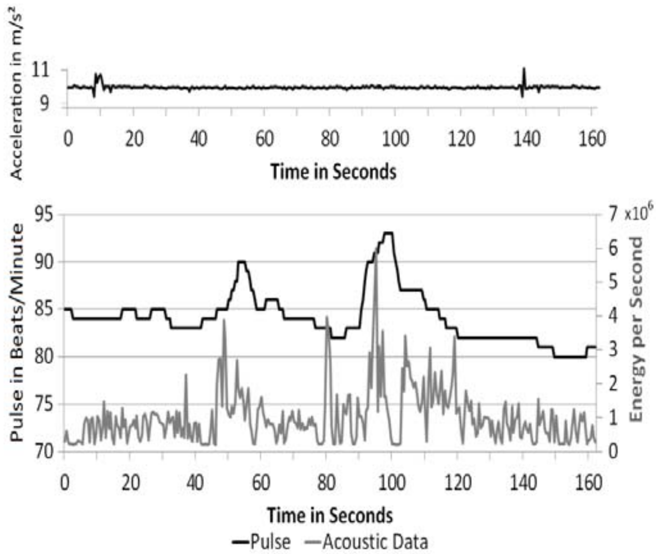

Figure 5: Pulse and acoustic signal while a proband is speaking with a supervisor.

4.1 Evaluation

Assuming correlations between the activity and stress levels of a person and environmental sound intensity in a work environment, we measured the person's load and sound intensity by use of a Smartphone worn in the trousers pocket of that person. As indicator for the stress level of the person the heart rate has been monitored.

As shown in figure 5, the first tests provide the indication that the heart rate rises when the surrounding sound increases even when the physical activity stays stable. Regarding sound and pulse rate, we can identify a correlation of sound energy and pulse rate of the test subject. In this scenario, the subject (a student) is talking easily to his supervisor (heart rate = 85 bpm). As soon as the subject proposes a statement, the heart rate rises to 93 bpm until the supervisor agrees to the statement. After this occurrence the pulse rate decreases to normal state. We assume that the stress is correlating to the pulse rate, so that in this example the sound correlates well with the perceived stress of the subject.

4.2 Discussion

We developed a novel method of activity recognition by the combination of sound data and acceleration patterns. For this purpose the DiaTrace activity recognition system has been extended by algorithms analyzing audio data. We could show that fusing both information sources improves the accuracy of physical activity recognition in real-world settings.

By using the accelerometer as well as the microphone of a Smartphone, not only several different bodily activities can be detected but also the intensity of the environmental noise. We determined that sound recognition is well possible with a Smartphone, even when the device is worn in a trousers pocket. Furthermore we analyzed the recorded audio data of the microphone and could differ between activities like driving a car, cycling, or working at office.

5. CONCLUSIONS AND OUTLOOK

This paper presents the new methodology of using both sound and acceleration sensor of a commercial smart phone for activity recognition. Microphones in mobile phones are suitable as sound sensors even when they are in the trousers pocket. The movement of the mobile devices causes a significant sound pattern as well as the ambient sound provides a significant sound pattern to retrieve more information about the current environment. The paper illustrates the differences between some activity classes and presented sound features to support the classification.

As the work presented here is in its initial phase, more evaluation tests have to be performed in order to find out about the general applicability of the algorithms applied on the sound data. Also we will investigate the applicability of the approach for settings in which the phone is not in the trousers pocket but a handbag or rucksack close by the person as is often the case with women.

Further work will consider identification of sound caused by other people and that of machinery. This might be significant for the assessment of social contacts, work environment, and context changes of a person. We expect that sound data analysis on mobile phones will also become relevant for estimation of working conditions (sound pollution) in e.g. schools or for recreation periods in breaks or at free time.

6. REFERENCES

[1] DeVaul R., Sung M, Gips J., Pentland A. , MIThril: Applications and Architecture, Proc. 7th IEEE International Symposium on Wearable Computers, White Planes, NY, USA, 21-23 October 2003.

[2] Bieber G., Voskamp J., Urban B., Activity Recognition for Everyday Life on Mobile Phones, Lecture Notes in Computer Science, Volume 5615/2009, pp.289-296, Springer Berlin / Heidelberg, 2009, ISSN 0302-9743

[3] Peter, C., Bieber G., Urban B. :Affect- and Behaviour-Related Assistance for Families in the Home Environment. The 3rd ACM International Conference on PErvasive Technologies Related to Assistive Environments : PETRA 2010 [CD-ROM].New York : ACM, 2010, ISBN 978-1-4503-0071-1

[4] Wichtendahl S., Denton R. Cost effective predictive maintenance solutions, Wilcoxon Research, Inc., 21 Firstfield Rd, Gaithersburg, MD 20878, USA. Company White Paper, June 04, 2007

[5] Tajadura-Jiménez A., Västfjäll D. : Auditory-induced emotion: a neglected channel in human-computer interaction. The 20th British HCI Group Annual Conference 11-15 Sept. 2006 Queen Mary, University of London.

[6] Wolfson S., Case G.: The effects of sound and colour on responses to a computer game. Interacting with Computers 13 (2000) 183-192, (c) 2000 Elsevier Science B.V.

[7] Rolls E., Emotion explained, Oxford University Press, 2005, ISBN-13: 978-0198570035

[8] Dai J., Bai X., Yang Z, Shen Z., Xuan D.: Mobile phone-based pervasive fall detection, Pers Ubiquit Comput, DOI 10.1007/s00779-010-0292-x, Springer Verlag London, 2010

[9] Paerkkae, J., Ermes M., Korpipaae P., Maentyjaervi J., Peltola J., Korhonen I.: Activity Classification Using Realistic Data From Wearable Sensors, IEEE Transaction on Information Technology in Biomedicine, Vol 10, No. 1, 2006

[10] Richter, Claas; Tanaka, Yuzuru; Bieber, Gerald; Gohlke, Holger:A Case Study of an Employee Portal with integrated Physical-Activity Monitoring.In: Cleve, Jürgen (Hrsg.); Hochschule Wismar:7. Wismarer Wirtschaftsinformatiktage : WIWITA 2010.Wismar, 2010, pp. 271-281

[11] Schiel, Ralf; Kaps, Alexander; Bieber, Gerald; Kramer, Guido; Seebach, Henning; Hoffmeyer, André:Identification of Determinants for Weight Reduction in Overweight and Obese Children and Adolescents.In: Journal of Telemedicine and Telecare 16 (2010), 7, pp. 368-373

[12] Chipchase J., Yanqing C., Ichikawa F., Where’s The Phone? Selected Data, survey, NOKIA, 2007