User Activity Related Data Sets for Context Recognition

Holger Junker[1], Paul Lukowicz[1,2], Gerhard Tröster[1]

[1]Wearable Computing Lab, ETH Zurich, Switzerland
[2]Institute for Computer Systems and Networks, UMIT Innsbruck, Austria
(junker, lukowicz, troester@ife.ee.ethz.ch)

Abstract. The use of body-worn sensors for recognizing a person's context has gained much popularity recently. For the development of suitable context recognition approaches and their evaluation, real-world data is essential. In this paper, we present two data sets which we recorded to evaluate the usefulness of sensors and to develop, test and improve our recognition strategies with respect to two specific recognition tasks.

1 Introduction

For the development and evaluation of recognition strategies real-world data is required. However, collecting such data can often be very time consuming slowing down the overall development process. In recent years, a lot of different recognition tasks have been investigated ranging from the recognition of user activities such as walking, ascending and descending stairs or cycling, up to the recognition of emotional states[1]. For the evaluation of feature extraction schemes, fusion strategies and classification algorithms, researchers have mostly generated and used proprietary data. Based on the fact that many researchers often focus on similar or even identical recognition tasks, the availability of common public access data sets would increase the possibility of having statistically relevant data, improve the quality of results and speed up the systematic development of methods. This has been acknowledged in various fields of research, e.g. in biometric authentication[2] where the development of methods is speeded up with competitions using public databases.

Contributing to those facts, we present two of our data sets that we recorded for evaluating specific recognition tasks which could be made available for public use. Both data sets are related to the domain of user activities. In the following we briefly describe those two data sets.

2 Data Set 1: Simple User Activities

Many publications dealing with the recognition of user activities have targeted the recognition of fairly simple activities such as sitting, standing, walking, ascending stairs and descending stairs. Recognizing such activities is just a first step towards systems that one day will allow the recognition of much more complex activities. Despite their simplicity, this simple scenario where subjects walk a predefined path including level walking, ascending and descending stairs have become a quasi-standard scenario that is often used to

[1] http://affect.media.mit.edu/
[2] http://bias.csr.unibo.it/fvc2004

Table 1. Sensing Modalities, Sensor Location and Features extracted

Sensing modalities	**Sensor location**
2 three-axis accelerometers 1 Air pressure sensor	1st accelerometer: above right knee, with axes aligned to antero-posterior and vertical body axis; 2nd accelerometer: Back of body, mounted on belt (same axes orientation as 1st accelerometer: Inside right, inner pocket of jacket
1 One-axis gyroscope	Above right knee, sensitive axis aligned with lateral axis
2 Force Sensitive Resistors (FSR)	1st FSR: mounted under heel of right shoe sole; 2nd FSR: mounted under ball of right shoe sole. Used to measure initial ground contact of heel and ball, respectively.

investigate issues such as the usefulness of specific sensors, specific classification algorithms and feature extraction schemes. Providing benchmark data sets based on this scenario will be very useful for researchers in the area of context recognition. In order to investigate the usefulness of different sensing modalities and their corresponding features for the envisioned recognition task, we equipped four test subjects with the sensors and instructed them to walk repeatedly a predefined path including stairway and level walking, without any further instructions, e.g. concerning speed of walking.

2.1 Sensors & Placement

Table 1 lists the sensors and placements used for the experiments.

The use of accelerometers and gyroscopes for classifying level walking and stairway walking have been investigated by many researchers [6,8,7,3]. Air pressure sensors have been evaluated in [8,5] and force sensitive resistors in [1]. Although the list of sensors used in our experiments is not exhaustive, it covers most of the sensors recently proposed for the recognition of the different modes of locomotion.

2.2 Labelling

Labelling was carried out manually during the experiment by the individual test subjects and recorded together with the sensor data providing labelling accuracy in the range of approximately 1 s (depending on the reaction time of the subject). Walking on landings between two stairs was labelled as level walking. Stairway walking started when a foot hit a step.

2.3 Data Format & Technicalities

Technicalities All sensors, except for the air pressure sensor (Intersema MS5534A), have been synchronously sampled with 100 Hz and converted with 12 Bit resolution[3]. All data is available in simple ASCII-File Format with the sensor readings stored in columns.

[3] The air pressure sensor was sampled with 1 Hz an 16 bit resolution

3 Data Set 2: Wood Workshop Tasks

As part of our work analyzing human activity using on body sensors, the Wood Workshop experiment provides a rich source of data for research into activity recognition using distributed microphones and accelerometers. The details of our initial experiment can be found in the proceedings of this conference [4].

3.1 Activity Data

In the work presented, data is collected from a single subject performing a sequence of assembly tasks in a wood workshop. More recent experiments involve up to five different subjects. The exact sequence of actions for all of these recordings is listed in Table 2.

No	action
1	take the wood out of the drawer
2	put the wood into the vise
3	take out the saw
4	saw
5	put the saw into the drawer
6	take the wood out of the vise
7	drill
8	get the nail and the hammer
9	hammer
10	put away the hammer, get the driver and the screw
11	drive the screw in
12	put away the driver
13	pick up the metal
14	grind
15	put away the metal, pick up the wood
16	put the wood into the vise
17	take the file out of the drawer
18	file
19	put away the file, take the sandpaper
20	sand
21	take the wood out of the vise

Table 2. Steps of workshop assembly task

The experiment was repeated 10 times in the same sequence to collect data for training and testing. For practical reasons, the individual processing steps were only executed long enough to obtain an adequate sample of the activity. This policy did not require the complete execution of any one task (e.g. the wood was not completely sawn), allowing us to complete the experiment in a reasonable amount of time. However this protocol influenced only the duration of each activity and not the manner in which it was performed.

3.2 Sensors & Placement

The data was collected using the ETH PadNET sensor network [2] equipped with 3 axis accelerometer nodes and two Sony mono microphones connected to a body worn computer.

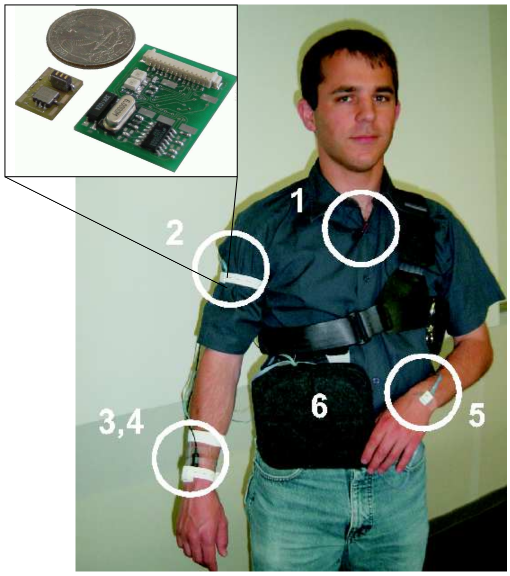

Fig. 1. The wood workshop (*left*) with *(1)* grinder, *(2)* drill, *(3)* file and saw, *(4)* vise, and *(5)* cabinet with drawers. The sensor type and placement (*right*): *(1,4)* microphone, *(2,3,5)* 3-axis acceleration sensors and *(6)* computer

The position of the sensors on the body is shown in Figure 1: an accelerometer node on both wrist and on the upper arm of the right hand, and a microphone on the chest and on the right wrist (the initial subject was right handed).

Audio data is stored in two channel (one for each mic) WAV format at a recording rate of 44kHz. For the recognition algorithms, this was downsampled to 4.8kHz. One WAV file is stored per recorded sequence.

The six continuous accelerometer feature values (wrist x,y,z; elbow x,y,z) are recorded as columns in an ASCII file, one for each sequence, at a rate of approximately 93 Hz.

3.3 Labelling

In order to perform supervised learning, and to be able to calculate recognition rates, the recorded data had to be segmented according to the activity classes that we wish to recognize.

The first problem is in deciding exactly what constitutes an activity that we could classify under a single label. In hammering, for example, should we aim to recognize only the stroke of the hammer hitting a nail, or should we include the entire hammering sequence? We choose, in the interests of simplicity, the second approach, defining activities as: hammering, sawing, filing, drilling, etc. These are all activities which, at least for these recordings, last more than a few seconds and may be regarded as, at least for sound, quasi-stationary for their duration.

In the first recordings, labelling was performed by observing and listening to the raw data, marking out segments of interest by hand. Given the repetitive nature of the sequence used, this approach allowed labelling to be carried out with a good degree of accuracy, but at the expense of being rather time consuming.

For larger data sets, this is impractical, and so a method of performing time-of-recording labelling was investigated. This involved a second person observing the subject, pressing keys on a terminal to mark the start and stop times of the various activities. Initially, this was purely an observational labelling, but errors by both subject and observer - in particular the often poor reaction time of the observer - led to development of a semi-autonomous command approach. With this scheme, the computer issues instructions, see Table 2, which the observer instructs the subject to perform; this gives the observer time to press the activity start/stop key, thus reducing the potential for timing errors. As the sequence should be in order, the only labelling errors occur when a subject performs an activity that they were not instructed to do. In such cases, the observer annotates the recording so that the labelling can be later amended by hand.

3.4 Synchronization of Multiple Data Types

In order to ensure that any recognition based on sensor data fusion will work, and that the labels are meaningful, the incoming sensor data must be appropriately synchronized. This is achieved by punctuating the beginning and end of every sequence with the easily distinguishable gesture, in both sound and acceleration, of clapping. So long as the clap appears towards the very beginning of a recording, a simple peak detection algorithm can be tuned to find corresponding points on both the audio and acceleration clap.

In the recognition experiments carried out on this data, the timescales of interest, e.g. the duration of a hammering activity, is in the order of seconds. For this reason, and the fact that the labelling itself is subject to human response errors of the same order, exact synchronization is not an issue.

References

1. H. Junker, P. Lukowicz, and G. Troester. Locomotion analysis using a simple feature derived from force sensitive resistors. In *Proceedings Second IASTED International Conference on Biomedical Engineering, 2004*, 2004.
2. N. Kern, B. Schiele, H. Junker, P. Lukowicz, and G. Tröster. Wearable sensing to annotate meeting recordings. In *6th Int'l Symposium on Wearable Computers*, pages 186–193, October 2002.
3. J. Mantyjarvi, J. Himberg, and T. Seppanen. Recognizing human motion with multiple acceleration sensors. In *2001 IEEE International Conference on Systems, Man and Cybernetics*, volume 3494, pages 747–752, 2001.
4. Holger Junker Gerhard Troster Paul Lukowicz, Jamie A Ward and Thad Starner Amin Atrash. Recognizing workshop activity using body worn microphones and accelerometers. In *Pervasive Computing*, 2004.
5. K. Sagawa, T. Ishihara, A. Ina, and H. Inooka. Classification of human moving patterns using air pressure and acceleration. In *IECON '98*, volume 2, 1998.
6. M. Sekine, T. Tamura, T. Fujimoto, and Y. Fukui. Classification of walking pattern using acceleration waveform in elderly people. In D. Enderle, J., editor, *Proceedings of the 22nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society Cat*, volume 2, 2000.

7. L. Seon-Woo and K. Mase. Recognition of walking behaviors for pedestrian navigation. In *Proceedings of the 2001 IEEE International Conference on Control Applications (CCA'01) (Cat*, pages 1152–1155, 2001.
8. K. Van-Laerhoven and O. Cakmakci. What shall we teach our pants? In *Digest of Papers. Fourth International Symposium on Wearable Computers.*, pages 77–83, 2000.