A Novel Method for the Automatic Segmentation of Activity Data from a Wrist Worn Device: Preliminary Results

James D. Amor *IEEE Member*, Vijayalakshmi Ahanathapillai, and Christopher J. James, *IEEE Senior Member*

Abstract—**Activity monitoring is used in a number of fields in order to assess the physical activity of the user. Applications include health and well-being, rehabilitation and enhancing independent living. Data are often gathered from multiple accelerometers and analysis focuses on multi-parametric classification. For longer term monitoring this is unsuitable and it is desirable to develop a method for the precise analysis of activity data with respect to time. This paper presents the initial results of a novel approach to this problem which is capable of segmenting activity data collected from a single accelerometer recording naturalized activity.**

I. Introduction

Activity monitoring is the practice of monitoring a user's physical activity. This can be done for a number of reasons, including health applications [1], lifestyle tracking [5] and sports performance measurement [3]. Activity monitoring can also take place over varying timescales from short, minutes-long monitoring, to longitudinal monitoring, lasting weeks or months. Activity monitoring is achieved principally through the use of accelerometers to monitor the movement of the subject and depending on the end goal and timeframe over which data are to be collected there are a variety of placement locations of the accelerometers. Placement on the wrist, trunk, thigh and waist are common. For longer term monitoring, an unobtrusive approach is often used and for this purpose placement on the wrist or waist is often preferred [1][4].

An activity monitor will record the user's activity using an accelerometer (typically in 3-axes) and provide this data for analysis. Traditional approaches have often focused on time domain analysis or multi-parametric classification, often from multiple devices [2][5]. Furthermore, these approaches are designed to classify activity into one of a set of predetermined categories. Whilst this approach works in a lab-based environment, extending it to the real world becomes problematic. Wearing multiple accelerometers is often not feasible for long periods and real world data rarely falls into neat categories anyway. Longer term monitoring also has a tendency to use coarsely grained analysis with respect to time due to the volume of data being analysed.

It is desirable therefore to provide a method capable of more precise analysis (fine time granularity) from a single device without simply breaking data into smaller packets. In this paper we take the approach of segmenting activity data such that each segment contains similar activity and that detected transition markers show where posture or activity has changed (without any attempt at actually classifying activity). We present a novel methodology to segment naturalised activity monitoring from a single device in this way as a precursor to further analysis on the detected segments. The results from a proof of principal study to determine if the method can be used to detect activity transitions in data from a wrist worn tri-axial accelerometer are also presented.

II. Experimental Method

In order to develop and test an algorithm to perform automatic segmentation we performed a data gathering experiment to obtain data from a simulated office environment. The aim of the experiment was to gather tri-axis accelerometer data and annotate this data in fine detail with respect to the actions of the participant so that the data-set can be used in a range of algorithm development tasks.

To this end participants were asked to perform a variety of tasks that simulated office work whilst wearing an

Figure 1. The FOV provided by the camera, showing the PC, the table and the office door.

Research supported by EU FP7 Reference Number 288532.

J. D. Amor was with the Institute of Digital Healthcare, WMG, University of Warwick, Coventry, UK and is now with the School of Engineering at the University of Warwick. (phone: +44 (0)24 76151939; e-mail: J.D.Amor@Warwick.ac.uk).

V. Ahanathapillai is with Institute of Digital Healthcare, WMG, University of Warwick, Coventry, UK. (email V.Ahanathapillai@Warwick.ac.uk)

C.J. James is with Warwick Engineering in Biomedicine, School of Engineering, University of Warwick, Coventry, UK. (email C.James@Warwick.ac.uk).

978-1-4244-7929-0/14/$26.00 ©2014 IEEE

accelerometer (an Actigraph GT3X+) to record raw data. The accelerometer was placed on the wrist. Each experimental session was videoed using a fixed camera and the sessions were manually annotated *post-hoc* by two researchers.

The activities the participants were asked to do are shown in Table 1 and each task was timed to approximately five minutes duration. The tasks are designed to mimic work in an office environment and aim to occupy the participant enough that they performed the task in as natural a way as possible. As such tasks were designed to engage the participant's attention.

TABLE I. PARTICIPANT INSTRUCTIONS FOR DATA GATHERING

Task	Name	Participant Instructions
1	Reading	Start outside office. Open door. Sit at table. Read magazine.
2	Make drink	Leave office; walk to kitchen. Make hot drink. Walk back to office. Sit at table. Conversation with researcher.
3	Read website	Get up and sit at PC. Open browser. Go to news website. Read news website.
4	Typing	Open MS Word. Type document[a] into Word. Format double line spaced.
5	Ammend Doc.	Print document. Collect from printer. Sit at table. Highlight every occurrence of the word "the". Make written amendments to document so that it would be suitable for a 5 year old to read.
6	Walking 1	Get up. Go for walk with researcher and return to office.
7	Lifting and Writing	Back in office. Lift box from shelf to table. Remove contents from box to table. Arrange in order of size. Place in a pile. Put back in box. Get puzzle book out and pencil. Sit. Do puzzles.
8	Watching Screen	Put puzzle book back. Put box on floor by PC. Sit at PC. Watch internet videos.
9	Reading and Writing	Pick up box. Go to table. Repeat exercise with box contents. Sit. Remove book. Look up topical[b] information. Write interesting facts and draw a diagram.
10	Walking 2	Repeat of Walking 1 in reverse direction.
11	Standing Idle	Stand idle, conversation with researcher.

a The document to be typed was a brief history of the University of Warwick.

b The topic was picked at random from the textbook each time.

A fixed position web-camera was used to record the experimental sessions and was positioned to enable the participant to remain in the field of view (FOV) of the camera throughout. The FOV of the camera can be seen in Figure 1. The exceptions to the subject remaining in the FOV of the camera were the activities that took place outside of the office. These are the trip to make a drink (activity 2) and the two walks (activities 6 and 10). As activity occurred outside of the FOV of the camera, detailed annotation cannot be maintained for these sections and as such they are excluded from analysis. These data will be used in other areas of activity monitoring research.

The videos of the sessions were used to produce fine grained annotations of the activity of the participant which act as a ground truth for the analysis of the accelerometery data. Annotations were made in accordance with Tables II and III. In addition to the timestamp, two annotations were recorded; the instantaneous activity and the prevailing activity. The instantaneous activity records point-events, which are those events that occupy a point in time (in practice a short period of time) such as postural transitions, position changes and miscellaneous movement. Prevailing activities, record the ongoing background posture and activity level onto which the point activities sit and comprise a *posture-activity* pairing.

TABLE II. PREVAILING ACTIVITY ANNOTATIONS

Annotation	Usage Notes
Crouching {inactive, semi-active, active})	The participant is crouched down, in some activity state.
Sitting {inactive, semi-active, active}	The participant is sitting in a chair, in some activity state.
Standing {inactive, semi-active, active}	The participant is standing, in some activity state.
Walk {short, long}	The participant is walking. A distance of short indicates walking around the office. A distance of long indicates walking outside the office.
Standing active and short walk	The participant is standing in an active state and moving up to 3 steps before satanding in an active state again.

TABLE III. INSTANTANEOUS ACTIVITY ANNOTATIONS

Annotation	Usage Notes
Transition {posture}-to-{posture}	Used when the participant changed postures (defined as crouching, sitting, standing and walking).
Activity Change	Used when the participant changed activity intensity (defined as inactive, semi-active or active) but not posture.
Position Change {minor, major}	Used when the participant changed position but not posture or activity intensity. The degree of movement during the change determines the minor or major nature of the change.
Miscelanious Movement	Used to note any movement not falling into one of the above categories.
Office Door {Open, Closed, Locked, Unlocked}	Used to note the action of the office door.
Unit {On, Off}	Use when the wrist unit was taken on or off.

Data were gathered from 4 participants and analyzed using the methodology in Section III to compare automatically detected transitions to the annotation points.

III. DATA ANALYSIS

The accelerometery data are analyzed to automatically extract transition points in the data. The data, in each axis, are pre-processed to remove device noise (micro variations in the baseline that derive from the hardware measurement accuracy) using the baseline smoothing method [1]. The data are then low pass filtered with an anti-aliasing filter and resampled from the collection frequency of 100 Hz down to 30 Hz. This initial step is performed to remove the influence of high frequency noise in the signal. Following on from this, a spectrogram is created using the Welch method [6] spectral estimate from an overlapping sliding window with length n_{win} and overlap percentage p_{over} on the data. This produces a 2D matrix, the columns of which correspond to time points and the rows of which correspond to frequency. From this matrix, a distance measure is calculated between adjacent columns using the Manhattan distance measure to create a vector of differences. An average value is computed over all the axes and this average difference vector is then thresholded; any differences above the threshold are marked up as a detected transition. The time of the detected transition is calculated as

$$t^m = 0.5\left(t_j^w + t_{j+1}^w\right) \tag{1}$$

where t^m is the transition marker time and t_x^w is the time of a window from the spectrogram, defined as

$$t_x^w = t_0 + 0.5(n_{win}) + (x-1)(n_{inc}), \tag{2}$$

where n_{inc} is the increment length in the sliding window operation defined as

$$n_{inc} = (n_{win} \times p_{over}) \tag{3}$$

A. Time resolution

From the description of the algorithm it can be seen that there is a time resolution that the algorithm functions at – that is, the exactness with which a change can be detected. The algorithm is sensitive to both the time window and overlap used when calculating the spectrogram; a larger window, or shorter overlap will serve to decrease the time resolution of the algorithm. The combination of these two factors serves to produce a time resolution, α, given by

$$\alpha = 2(n_{win}) - (p_{over} \times n_{win}). \tag{4}$$

B. Accuracy Assessment

To determine the accuracy of the algorithm an analysis with a leave one out approach is used; the threshold is trained on 3 data-sets and accuracy is tested on the remaining one. This is repeated for each data-set. A window size of 10 seconds with a 0.75 overlap was used as these were empirically determined to be the best settings.

The accuracy of the detector is assessed by comparing the output markers with the annotation list and identifying true positives (a detected marker with a matching annotation), false positives (a detected marker with no matching annotation) and false negatives (an annotation with no matching detected markers). A match between an annotation and a detected marker is recorded when the detected marker falls within range of the annotation. This range, which accounts for the time resolution, is defined as

$$\beta = 0.5(\alpha) + 2(\tau), \tag{5}$$

where τ is set to 2 seconds and accounts for any inaccuracies in the manual annotation timing made by the researchers. Accuracy results are given in terms of precision (proportion of markers that are correct, p), recall (proportion of annotations returned, r) and F_1 score, which is a weighted combination of precision and recall defined as

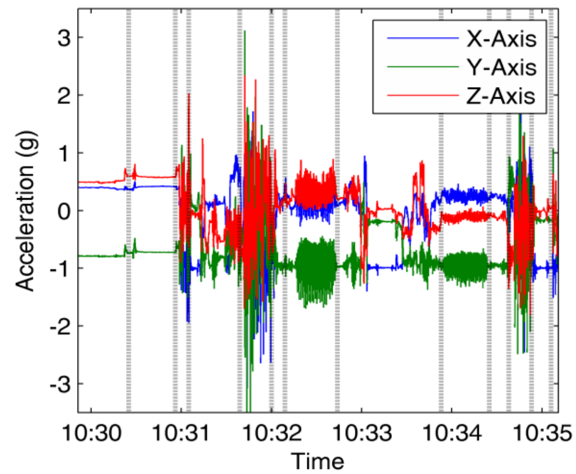

Figure 3. Raw Data and Annotated Transitions. Black vertical lines show annotation positions.

$$F_1 = 2 \times \frac{(p \times r)}{p + r}. \tag{6}$$

The F_1 score is used in this case because it weights precision and recall equally and aims to produce an output with a balance of the two features.

IV. Results

Figure 2 shows an example of the raw data obtained from the trial and the annotation points that were marked up in the data. A visual examination of the data shows that the annotation points lie on transition points in the accelerometery data. This is mirrored in Figure 3, which shows a spectrogram of the same data (z-axis). A visual examination of the spectrogram shows that there are blocks of similar looking activity with transition points between them. The transition points in the spectrogram match up with changes in the raw data signal and also map to some extent to the manually identified annotations. These transition points

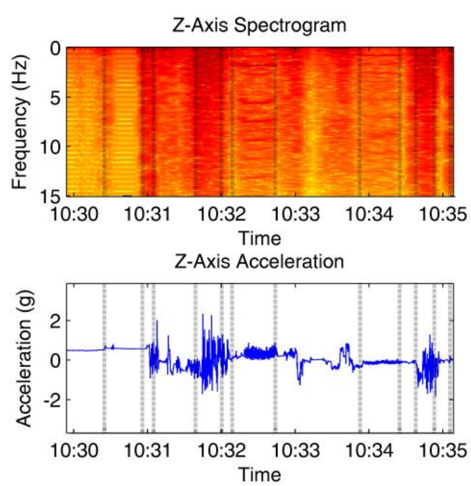

Figure 2. Spectrogram of the Z-Axis data showing the spectrum and the raw data. Black vertical lines show annotation positions.

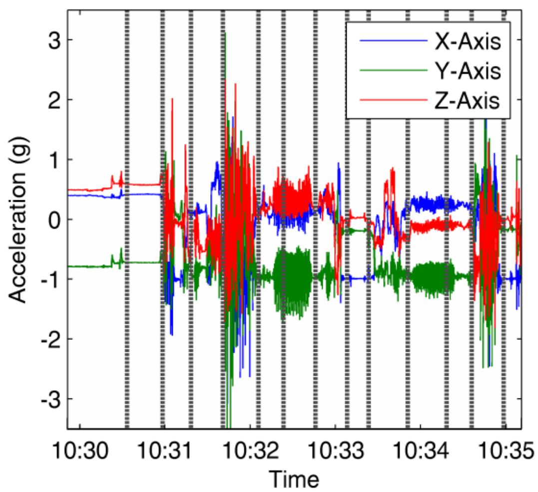

Figure 4. Raw data and detected transitions. Vertical dashed lines inicate detected transition points.

are the target of the segmentation algorithm.

Figure 4 shows an example of detected markers overlaid on the raw data. A visual inspection of this data shows that, as with the annotated data, the detected markers tend to fall on transition points in the data.

Accuracy results for a test window length of 10 seconds with an overlap of 0.75 are shown in Table IV. The table shows the F_1 scores along with the precision and recall scores for each data-set. It can be seen from the table that recall is significantly higher than precision across the board.

TABLE IV. ACCURACY RESULTS FOR 10 SECOND WINDOW

Test Data	F_1	*Precision*	*Recall*
1	0.5096	0.4255	0.6349
2	0.5644	0.4318	0.8143
3	0.5426	0.4544	0.6711
4	0.4276	0.3131	0.6739
Avg.	0.51105	0.4062	0.69855

V. DISCUSSION

The results obtained show that this method is capable of detecting the annotated transition points in the data with an accuracy of around 51%, when measured with an F_1 score, at a time resolution of 15 seconds (using a 10 second window with a 0.75 overlap). The results also show that recall is significantly higher than precision across the board which means that a large proportion of the annotated points are being detected. There is a trade off in that this produces a number of false positives, but it is far easier to remove false positives than to re-discover false negatives and this is one avenue that will be explored further.

It should also be noted that the results could be affected by the quality and consistency of annotation in the first instance. Incorrect annotation will be reflected in poor results performance. Whilst care has been taken to minimize this error, some element will remain. A further avenue of research would look manually at each detected marker and analyze the video to determine the validity of the algorithmic judgment.

The accuracy score, whilst relatively poor does not reflect fully the performance of the segmentation algorithm. When the data is examined at a coarse scale (over the entire length of the experiment) it can be seen that, whilst the specific timings identified by the algorithm and the annotations do not match up very well, the general areas where annotations occur and markers are detected to match to a large degree. This indicates that the algorithm is identifying periods of disturbance and not falsely identifying periods of relative stability in terms of the actions of the participant.

There is significant further work that can be performed to improve the detector. Different distance measures could be used and these might improve performance. Of particular interest is the cepstral distance, which is used in speech processing. Additional work may also focus on post processing of identified markers to reduce the false positive rate and an in depth analysis of the markers and the ground truth video could be performed to identify the validity of the automated detection. Finally, an analysis of the performance of the algorithm against others in the field and on different datasets could be undertaken.

VI. CONCLUSIONS

The work presented here has shown initial steps towards a novel automatic segmentation algorithm for human activity monitoring that is capable of segmenting naturalized data collected from a wrist worn accelerometer. The results show that the detector performs reasonably on the data and that there are several avenues of further research available to refine the detector performance. Based on the results obtained we consider the approach to be worthy of further work and consideration.

ACKNOWLEDGMENT

The authors would like thank the members of the EU FP7 USEFIL project (reference number: 288532) for their work and support.

REFERENCES

[1] J.D. Amor, V. Ahanathapillai1, and C.J. James. “Detecting and Analyzing Activity Levels for the Wrist Wearable Unit in the USEFIL Project.” *XIII Mediterranean Conference on Medical and Biological Engineering and Computing 2013 IFMBE Proceedings*. Volume 41, 2014, pp 1201-1204.

[2] L. Bao and S.S. Intille. “Activity Recognition from User-Annotated Acceleration Data” A. Ferscha and F. Mattern (Eds.): PERVASIVE 2004, LNCS 3001, pp. 1–17, 2004.

[3] James, D.A. “An accelerometer based sensor platform for insitu elite athlete performance analysis.” Sensors, 2004. Proceedings of IEEE. 1373 - 1376 vol.3. 2004.

[4] M.J. Mathie, A.C.F. Coster, N.H. Lovell, and B.G. Celler. "Accelerometry: providing an integrated, practical method for long-term, ambulatory monitoring of human movement." *Physiological measurement* 25, no. 2 (2004): R1.

[5] J. Parkka, E. Miikka, K. Panu, M. Jani, P. Johannes, and I Korhonen. "Activity classification using realistic data from wearable sensors." *Information Technology in Biomedicine, IEEE Transactions on* 10, no. 1 (2006): 119-128.

[6] P.D. Welch. "The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms." *IEEE Transactions on audio and electroacoustics* 15.2 (1967): 70-73.