Evaluating the effects of signal segmentation on activity recognition

Oresti Banos*, Juan-Manuel Galvez, Miguel Damas, Alberto Guillen, Luis-Javier Herrera, Hector Pomares, and Ignacio Rojas

Department of Computer Architecture and Computer Technology, Research Center for Information and Communications Technologies - University of Granada (CITIC-UGR)
C/Periodista Daniel Saucedo Aranda s/n, 18071 Granada, Spain
{*oresti,jonas,mdamas,aguillen,jherrera,hector,irojas*}*@ugr.es*

Abstract. On-body activity recognition systems are becoming more and more frequent in people's lives. These systems normally register body motion signals through small sensors that are placed on the user. To perform the activity detection the signals must be adequately partitioned, however no clear consensus exists on how this should be done. More specifically, considered the sliding window technique the most widely used approach for segmentation, it is unclear which window size must be applied. This paper investigates the effects of the windowing procedure on the activity recognition process. To that end, diverse recognition systems are tested for several window sizes also including the figures used in previous works. From the study it may be concluded that reduced window sizes lead to a better recognition of the activities, which goes against the generalized idea of using long data windows.

Keywords: Activity recognition, Wearable sensors, Inertial sensing, Segmentation, Window size

1 Introduction

Human behavior inference has been widely explored during the last years. People activity inference or recognition is normally performed through the analysis of the body motion, for which inertial sensors placed on limbs and trunk are particularly used. Despite activity recognition has been traditionally restricted to the scope of research and in-lab experiments, an strong effort is being put to commercially leverage all the knowledge gained so far. In fact, several new gadgets and systems are released day-by-day and put at the reach of most people. Their applications, mostly dedicated to specific wellness domains, range from assessment of training routines [1], calculation of energy expenditure [7, 11] or evaluation of dietary habits [12]. Yet, these systems are far from being robust and accurate enough for a lifelong and extensive use.

* Corresponding author.

For the sake of recognition, the signals registered through the on-body sensors are processed. One of the most important stages of this processing is data segmentation. Although diverse segmentation approaches has been proposed in the past, the most extensively used for its implementational simplicity and potential use in real-time applications is the sliding window method. Here the signals are split in windows of a fixed size and with no inter-window gaps. An overlapping between adjacent windows is tolerated for certain applications, however this is less frequently used. A range of window sizes have been used in previous studies from 0.1 seconds [20] to 12.8 seconds [15] or more [25], with some studies including a degree of overlap between windows [6, 21, 16]. In most cases, these values are randomly selected and employed without considering the special needs of systems for a realistic setting. Actually, depending on the addressed problem a fast identification may be needed (e.g., fall detection) or conversely not have special time requirements (e.g., kilometers walked in a day). Since reducing the recognition time (i.e., segmentation) may have an influence on the system performance, a tradeoff between detection time and accuracy should be considered by recognition system designers. Despite the importance of this, little work has been devoted to investigate this fact.

In this work we present an extensive study of the effects of segmentation for diverse recognition techniques and activities. The performance of several recognition systems is evaluated for an extensive set of window sizes that also covers the values used in previous works. This characterization is defined for a wide variety of representative activities. The rest of the paper is structured as follows. In Section 2 the activity recognition methodology used in this study is described. Section 3 presents the results obtained for the different experiments while these are discussed in Section 4. Final conclusions and future work are shown in Section 5.

2 Activity Recognition Methods

Signal segmentation is one of the stages of the activity recognition process, also known as activity recognition chain. For on-body inertial sensing, raw unprocessed signals (normally acceleration) are collected through a set of sensors attached to the subject's body. Electronic noise or other kind of artifacts may disturb the measurements, thus sometimes a filtering process is applied [19]. Nevertheless, this is not always used since it may imply a certain information loss. In order to capture the dynamics of the signals these are divided into portions of data (i.e., segmentation). Then, a feature extraction process is performed to provide a handler representation of the signals for the pattern recognition stage. A wide range of heuristics [17], time/frequency domain [18] and other sophisticated mathematical and statistic functions [3] are commonly used. The feature vector is provided as input of a classifier or reasoner, which eventually provides the recognized activity. For multisensor configurations, decision aggregation or fusion could be also applied [4].

3 Results

3.1 Experimental setup

One of the most complete available activity recognition datasets [2] is here used for evaluating the effects of signal segmentation. This dataset comprises motion data recorded for 17 volunteers while performing 33 fitness activities. A set of nine inertial sensors attached to different parts of their bodies was used for the motion records. The potential of this dataset stems from the number of considered activities, diversity of body parts involved, as well as the variety in intensity and dynamicity of the actions. The use of multiple sensors permits to measure the motion (namely, acceleration, rate of turn and magnetic field orientation) experienced by each body limb and trunk, thus better capturing the body dynamics. From all recorded inertial magnitudes here only the acceleration data is considered since this demonstrates as the most prevalent sensor modality in previous activity recognition contributions. All the recordings were collected in an out-of-lab environment with no constraints on the way the activities must be executed.

The implemented recognition methods (Section 2) are now described. No preprocessing of the data is applied to avoid the removal of relevant information. This is normal practice when the activities are diverse. The segmentation process basically consists in a non-overlapping sliding window approach. Different window sizes are used for evaluation, concretely ranging from 0.25s to 7s in steps of 0.25s. This interval comprises most of the values used in previous activity recognition systems. Three feature sets (FS) are respectively used for evaluation: FS1='mean', FS2='mean and standard deviation' and FS3='mean, standard deviation, maximum, minimum and mean crossing rate'. These are features widely used in activity recognition [22, 10, 13] for their discrimination potential and ease of interpretation in the acceleration domain. Likewise, four of the most extensively and successfully machine learning techniques used in previous activity recognition problems are considered for classification: C4.5 decision trees (DT, [9]), k-nearest neighbors (KNN, [8]), naive Bayes (NB, [26]) and nearest centroid classifier (NCC, [14]). The k-value for the KNN model is particularly set to three. The evaluation of the activity recognition models is performed through a ten-fold random-partitioning cross validation process applied across all subjects and activities. The process is repeated 100 times for each method to ensure statistical robustness. To avoid data imbalancement artifacts, the *F-score* or F_1-*score* metric [24] is used to evaluate the performance of the recognition systems.

3.2 Window size evaluation

Figure 1 depicts the performance results obtained for each recognition method and for diverse window sizes. Similar tendencies are found for the performance of each individual classification technique for all feature sets. This determines that these results could be in principle generalized to other recognition models of similar nature. Systems based on FS3 (richest feature set of considered) provide better performance than for FS2, which also improve the results obtained for FS1.

The window size has a different performance impact for each classification paradigm. The performance of NB and NCC models increases as the size of the window grows. A minimum performance is obtained for 0.25s, which nevertheless increases up to 30% when the window is enlarged to 1 second. Actually, a 'cut-off' window size is found at 1 second for all feature sets. From that value on no significant benefits are obtained in general. For NB-FS1, less than 5% improvement is achieved for some random window sizes when compared to the performance at 1 second. This also applies to a lesser extent for the NB-FS2 model. Conversely, increasing the window size more than 2 seconds entails a worsening of the recognition performance for NCC-FS3. DT shows a top performance for window sizes between 1 and 2 seconds. Upper and lower values to these generally decrease the performance of the recognizer. The KNN model outstands among evaluated and allows us to maximally reduce the window size. This technique provides the highest performance, with an F_1-*score* above 0.95 for the simplest realization (FS1) and close to 1 for FS2 and FS3, all for minimum window sizes (0.25s-0.5s). For window sizes higher than 2 seconds for FS1 and FS3, and 3 seconds for FS2, the performance of the KNN systems decreases monotonically. The lowest performance is achieved for a window size of 7 seconds, which for some cases is up to 15% less than the baseline.

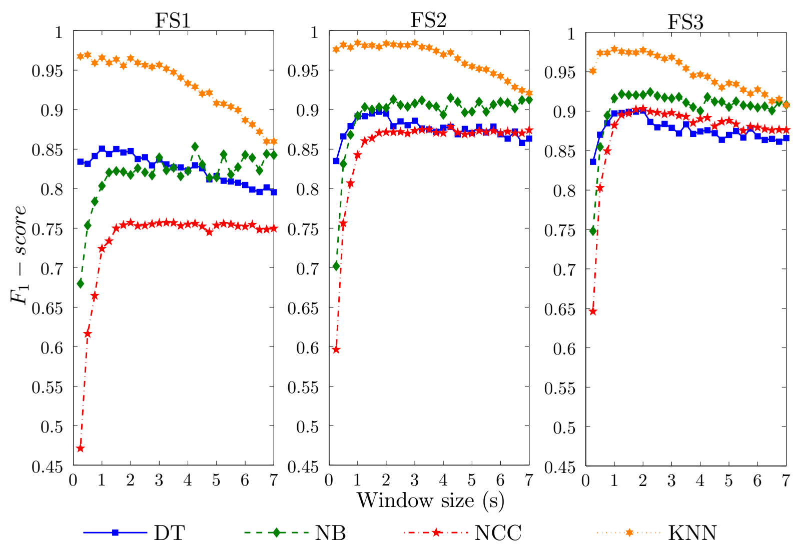

Fig. 1. Data window size effects on the activity recognition systems performance (F_1-*score*). Twelve recognition systems respectively corresponding to the combination of three feature sets (FS1, FS2, FS3) and four classification models (DT, NB, NCC, KNN) are evaluated.

4 Discussion

From the results obtained in Section 3.2 the interval 1-2 seconds demonstrates as the most reliable windowing range for DT and KNN models. Actually, KNN also provides very accurate recognition for minimum window size values (i.e., 0.25s-0.5s). NB and NCC methods, which provide the worst performance from evaluated, may benefit from adding more data for some sparse cases, however this may also lead to a performance drop. In this regard, the obtained results help reject the generalized idea of considering that the more data used for the feature extraction the most accurate the recognizer is. The diversity among activities determine this is just restricted to those of long duration or complex description, however many others benefit from shorter window sizes.

The results provided along with this work could be used as support tool for the design of activity recognition systems. When designing an activity recognition system, the expert may need to prioritize detection performance or speed, or even both. In most cases, a trade-off between both elements is required. As demonstrated in this study, in many cases a negligible reduction on the systems performance allow us to significantly shorten the window size. Systems that may benefit from rapid detections could not do it if a random window size is used. Moreover, other activities are better recognized for shorter window sizes. Therefore, the intuitive use of large windows could go in practice against the idea of optimizing the recognition capabilities.

Yet, there are some challenges and limitations that must be bore in mind for study. The presented results has been provided just for acceleration data, however current tendencies show that the use of other sensing modalities could help to improve recognition performance and systems robustness. Gyroscopes and magnetometers are more and more frequently used in combination with accelerometers for recognition purposes. Although accelerometers has been demonstrated to suffice, an analysis with these other modalities could be of interest. Moreover, a similar study to this could be also valuable for other activity recognition domains, such as for computer vision or ambient intelligence.

According to the considered recognition setup, to monitor several body parts multisensor configurations are required. Therefore, the results presented here are of limited application to those systems that rely on a very reduced set of sensors or even a unique device. Nevertheless, latest contributions show that ensuring robustness and guaranteeing a reasonable recognition rate demands a complete monitoring of the body as much as the number of target activities and their diversity increases [5, 23]. Thereby, we consider this study perfectly suits with current and specially future trends.

5 Conclusion

Signal segmentation is one of the main stages in the activity recognition chain. This process consists in the partitioning of the sensor data stream into smaller segments or windows. Most recognition systems use random window size values, however this may not optimally apply to the particular considered problem. Thus, a study that analyzes this is lacking.

This work presents an extensive study that analyzes the effects of the windowing process on the recognition systems performance. Several methodologies extensively used in previous works are considered for evaluation. From the results, short windows (2 seconds or less) demonstrate to provide the most accurate detection performance. In fact, the most precise recognizer is obtained for very short windows (0.25 - 0.5 seconds), thus proving that large window sizes not necessarily translates into a better recognition performance.

Systems configuration and design tasks may benefit from the figures provided as part of this work. Next steps include to extend the scope of this study to other activity recognition domains and technologies.

Acknowledgments. This work was partially supported by the Spanish CICYT Project SAF2010-20558, Junta de Andalucia Project P09-TIC-175476 and the FPU Spanish grant AP2009-2244.

References

1. Adidas®. Adidas micoach, 2013. http://micoach.adidas.com/.
2. Oresti Banos, Miguel Damas, Hctor Pomares, Ignacio Rojas, Mt Attila Toth, and Oliver Amft. A benchmark dataset to evaluate sensor displacement in activity recognition. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, pages 1026–1035, 2012.
3. Oresti Banos, Miguel Damas, Hector Pomares, Alberto Prieto, and Ignacio Rojas. Daily living activity recognition based on statistical feature quality group selection. *Expert Systems with Applications*, 39(9):8013 – 8021, 2012.
4. Oresti Banos, Miguel Damas, Hector Pomares, Fernando Rojas, Blanca Delgado-Marquez, and Olga Valenzuela. Human activity recognition based on a sensor weighting hierarchical classifier. *Soft Computing*, 17:333–343, 2013.
5. Oresti Banos, Miguel Damas, Hector Pomares, and Ignacio Rojas. On the use of sensor fusion to reduce the impact of rotational and additive noise in human activity recognition. *Sensors*, 12(6):8039–8054, 2012.
6. Ling Bao and Stephen S. Intille. Activity recognition from user-annotated acceleration data. In *Pervasive Computing*, volume 23, pages 1–17, 2004.
7. BodyMedia®. Bodymedia fit, 2013. http://www.bodymedia.com/.
8. T. Cover and P. Hart. Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13(1):21–27, january 1967.
9. Richard O. Duda, Peter E. Hart, and David G. Stork. *Pattern Classification (2nd Edition)*. Wiley-Interscience, 2000.
10. Davide Figo, Pedro C Diniz, Diogo R Ferreira, and João MP Cardoso. Preprocessing techniques for context recognition from accelerometer data. *Personal and Ubiquitous Computing*, 14(7):645–662, 2010.
11. Fitbit®. Fitbit products, 2013. http://www.fitbit.com/es/store.
12. Jawbone®. Jawbone up, 2013. https://jawbone.com/up/international.
13. Jennifer R Kwapisz, Gary M Weiss, and Samuel A Moore. Activity recognition using cell phone accelerometers. *17th Conference on Knowledge Discovery and Data Mining*, 12(2):74–82, 2011.

14. Wai Lam, Chi-Kin Keung, and Charles X. Ling. Learning good prototypes for classification using filtering and abstraction of instances. *Pattern Recognition*, 35(7):1491 – 1506, 2002.
15. Andrea Mannini, Stephen S Intille, Mary Rosenberger, Angelo M Sabatini, and William Haskell. Activity recognition using a single accelerometer placed at the wrist or ankle. *Medicine and science in sports and exercise*, 45(11):2193–2203, 2013.
16. Robin Marx. Ad-hoc accelerometer activity recognition in the iball. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, 2010.
17. Merryn J Mathie, Adelle C F Coster, Nigel H Lovell, and Branko G Celler. Accelerometry: providing an integrated, practical method for long-term, ambulatory monitoring of human movement. *Physiological Measurement*, 25(2):1–20, apr 2004.
18. U. Maurer, A. Smailagic, D.P. Siewiorek, and M. Deisher. Activity recognition and monitoring using multiple sensors on different body positions. In *International Workshop on Wearable and Implantable Body Sensor Networks*, pages 113–116, 2006.
19. B. Najafi, K. Aminian, A. Paraschiv-Ionescu, F. Loew, C. J. Bula, and P. Robert. Ambulatory system for human motion analysis using a kinematic sensor: Monitoring of daily physical activity in the elderly. *IEEE Transactions on Biomedical Engineering*, 50(6):711–723, 2003.
20. S. Pirttikangas, K. Fujinami, and T. Seppanen. Feature selection and activity recognition from wearable sensors. In *Third International Symposium on Ubiquitous Computing Systems*, volume 4239, pages 516–527, 2006.
21. Stephen J Preece, John Yannis Goulermas, Laurence PJ Kenney, and David Howard. A comparison of feature extraction methods for the classification of dynamic activities from accelerometer data. *IEEE Transactions on Biomedical Engineering*, 56(3):871–879, 2009.
22. Nishkam Ravi, Preetham Mysore, and Michael L. Littman. Activity recognition from accelerometer data. In *Seventeenth Conference on Innovative Applications of Artificial Intelligence*, pages 1541–1546, 2005.
23. Hesam Sagha, Hamidreza Bayati, Jose del R. Millan, and Ricardo Chavarriaga. Online anomaly detection and resilience in classifier ensembles. *Pattern Recognition Letters*, 34(15):1916 – 1927, 2013.
24. Marina Sokolova and Guy Lapalme. A systematic analysis of performance measures for classification tasks. *Information Processing & Management*, 45(4):427 – 437, 2009.
25. Maja Stikic, Tâm Huynh, Kristof Van Laerhoven, and Bernt Schiele. Adl recognition based on the combination of rfid and accelerometer sensing. In *Second International Conference on Pervasive Computing Technologies for Healthcare*, pages 258–263. IEEE, 2008.
26. Sergios Theodoridis and Konstantinos Koutroumbas. *Pattern Recognition*. Academic Press, 4th edition, 2008.