Novel Method for Feature-Set Ranking Applied to Physical Activity Recognition

Oresti Baños, Héctor Pomares, and Ignacio Rojas

Department of Computer Architecture and Computer Technology, University of Granada,
C/Periodista Daniel Saucedo Aranda s/n E-18071, Granada, Spain
oresti@correo.ugr.es, {hector,irojas}@ugr.es

Abstract. Considerable attention is recently being paid in e-health and e-monitoring to the recognition of motion, postures and physical exercises from signal activity analysis. Most works are based on knowledge extraction using features which permit to make decisions about the activity realized, being feature selection the most critical stage. Feature selection procedures based on wrapper methods or 'branch and bound' are highly computationally expensive. In this paper, we propose an alternative filter method using a feature-set ranking via a couple of two statistical criteria, which achieves remarkable accuracy rates in the classification process.

Keywords: Activity Recognition, Feature Selection, Ranking.

1 Introduction

Daily physical activity recognition is a very important task that is currently being applied in several fields such as e-health, robotics, sports, videogames industry, among others [4,7,9,10,11]. The primary difficulty consists in designing a system whose reliability is independent of who is carrying out the exercise and the particular activity execution style. Besides, the complexity is increased by distortion elements related to system monitoring and processing, along with the random character of the execution.

In the literature, most studies performed are based on supervised laboratory data. Nevertheless, the apparently good recognition results on supervised data that some works achieve cannot be extrapolated to unsupervised (semi-naturalistic) data [2,12]. In this paper we propose an automatic methodology to extract a set of the most important features to be used in activity recognition. One of the most important characteristic of the method proposed is that we do not provide a rank order for every individual feature but for every set of features, allowing for the synergical utility of several features when considered together at the same time.

The rest of the paper is organized as follows: In Section 2 we make a brief summary of the activity recognition process. Next, we present the rank-based feature-set selection methodology developed, describing the fundamentals of this method and the algorithm's main steps. Finally we evaluate the performance of the method for a specific example and we compare the accuracy results with related previous works.

N. García-Pedrajas et al. (Eds.): IEA/AIE 2010, Part II, LNAI 6097, pp. 637–642, 2010.

2 Activity Recognition

Our experimental setup starts from a signal set corresponding to acceleration values measured by a group of sensors located in strategic different parts of the body (hip, wrist, arm, ankle, thigh), for several daily activities (see Figure 1). This philosophy of work can easily be generalized to other studies related to activity recognition from a set of features.

The initial information provided by the sensors has some artefacts and noise associated to the acquisition data process. A low pass and high pass filtering is normally used to remove these irregularities.

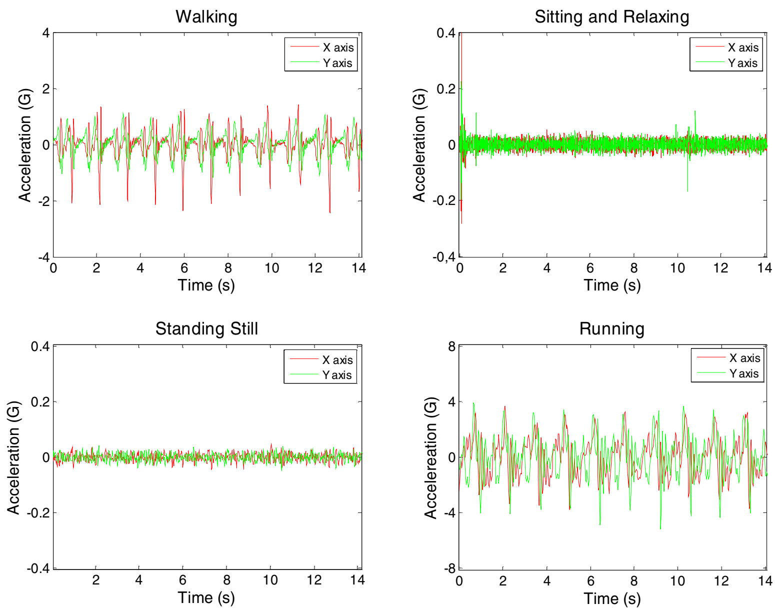

Fig. 1. Signals corresponding to 4 usual daily physical activities. It's relevant the signal pattern form similarity between "walking" and "running", and "sitting and relaxing" and "standing still" respectively.

Subsequently we generate a parameter set made up of 861 features corresponding to a combination of statistical functions such as mean, kurtosis, mode, variance, etc., and magnitudes obtained from a domain transformation of the original data such as energy spectral density, spectral coherence or wavelet decomposition, among others. These features are evaluated over the complete signal, although other alternatives based on windowing and sub-segmentation signal feature extraction could also been tested.

In this stage, we now must rely on a feature selection process that has the responsibility of deciding which features or magnitudes are the most important ones to decide the kind of activity the person is carrying out. In the next section we describe the method we have designed to accomplish this task.

3 Proposed Ranking Technique Based on Discrimination and Robustness

Obtaining a specific group of variables from a big initial set is not a trivial task, because the number of possible feature combinations is huge. In our experimental setup the sample space is represented by n = 861 features, so brute force techniques like 'branch and bound' ($O(2^n)$ convergence → $2^{861} \approx 1.5 \times 10^{259}$ possible permutations) or wrapper methods are impractical. In this section, we present an alternative method based on the concepts of discrimination and robustness for a complete set of features.

Let us define the *sample range* of a class as the set of values included between the maximum and the minimum value (both inclusive) that a feature or variable takes for this class. Every circle showed in the example of figure 2 represents a sample corresponding to the feature value calculated over the data from a concrete subject and class. We represent for every class the sample range using a right brace sign with the same class tonality.

Given a group of samples (associated to every class) we rank its discriminant capability with respect to that class through the overlapping probability between this class and the others. This is calculated computing the number of samples from the analyzed class which are inside of the sample range defined by the others. For N classes and M samples for each class (let us suppose that this number is independent of the class), we define the overlapping probability of a set of samples as follows:

$$p(k) = \frac{1}{N} \sum_{n \neq k}^{N} \frac{m(k,n)}{M} \tag{1}$$

with *m(k,n)* being the number of samples from the class *k* inside the sample range of class *n*.

In order to make this more understandable, for the example given in figure 2 (N = 4, M = 8), the overlapping probability for the class "running" is p = ¼(3/8+0/8+0/8) = 0.094, since there are 3 samples from the class "running" in the data range defined for class "walking", and 0 in the rest of classes. Consequently, this feature permits to discriminate a priori the activity *running* from the activity *standing still* or *sitting and relaxing*, but it could be mistaken with an approximately 9% probability with *walking*.

We now carry out a thresholding process which allows us to define the feature analyzed as discriminative or not. This *overlapping* threshold takes values from 0 (the most restrictive, for cases with no overlapping between classes) to 1 (the most relaxed, when every sample from a class is inside the others). In general for a specific feature, if the analyzed class exceeds the threshold, the feature will be considered as no discriminant for this class.

Apart from the discriminant capacity of a feature or a set of features, a second characteristic is now defined which takes into account the usability of this set of features in different information contexts or sources. For instance, a specific measure taken from, let's say, the ankle accelerometer can be very discriminative to distinguish between the activities *walking* and *standing still*, but this very same measure may not be that reliable when taken from the thigh accelerometer. There may be some measures with the same discriminant capability between those activities which are not so dependent of the exact location of the sensor or, at least, which are still reliable when taken from a bigger number of sensors. We will denote this measure as the robustness criterion of a set of features. In short, discriminant capacity says how useful a motion feature is in general, and robustness is how this depends on where the sensor is.

Combining both criteria we obtain a quality ranking procedure capable of grouping features in different stages. For the sake of simplicity, let us suppose a recognition system with 4 classes and 5 sources; features will be classified in groups defining a ranking (see table 1). For instance, features that discriminate 4 classes in every source will be added to group #1 (the best). Group #14 will be completed with features that classify 2 classes (the same) in 3 sources at least. This example is extensible to any classes and sources.

4 Results

To evaluate the effectiveness of the ranking method developed, we use a signal database[1] corresponding to the data monitored by 5 biaxial accelerometers (hip, wrist, arm, ankle and thigh) for 4 activities (introduced in figure 1) in laboratory (supervised) and semi-naturalistic (unsupervised) environments.

Most remarkable features (set #1 and #2 primarily) for supervised and unsupervised data are geometric mean for amplitude signal, autocorrelation and some wavelets coefficients obtained through a 5-level Daubechies decomposition. This together with a classification strategy based on C4.5 decision tree permits to achieve an accuracy rate close to 99% (98.92% ± 1.08%) for laboratory data and 95% (95.05% ± 1.20%) for semi-naturalistic data. We used a cross validation method for training and

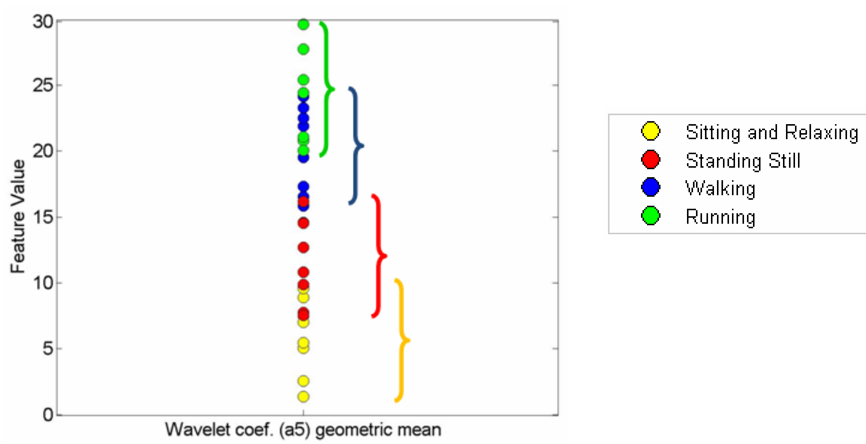

Fig. 2. Feature values extracted for the ankle accelerometer. Every circle represents the value (sample) of the feature "decomposition wavelet coefficients a5 geometric mean" for a specific subject (8 in all), identifying with different tonalities the belonging class.

[1] Database facilitated in [2] by Prof. Stephen Intille (Massachussets Institute of Technology).

Table 1. Example for 4 classes and 5 sources of quality feature set (ranking) based on discriminant (number of activities discriminated, first column) and robustness (number of motion sensor where the feature is discriminant, second column) criterions

Discriminant capacity	Robustness	Quality group
4	5	#1
4	4	#2
4	3	#3
4	2	#4
4	1	#5
3	5	#6
3	4	#7
3	3	#8
3	2	#9
3	1	#10
2	5	#11
2	4	#12
2	3	#13
2	2	#14
2	1	#15
1	5	#16
1	4	#17
1	3	#18
1	2	#19
1	1	#20
0	-	#21

testing. These results are quite good due to the relative parallelism between probability model used in feature selection defined and the entropy-based model used in decision tree.

Although a strict comparison with other studies cannot be made since the data and the number of classes may differ, in [9] a 83-90% classification accuracy was reached for laboratory conditions, 92.85%-95.91% in [8] (also for lab conditions), 89% in [2] for supervised and unsupervised, or 93% and 89% on recent works ([3] and [5] respectively) for semi-naturalistic data.

5 Conclusions

In this work we have very briefly shown a direct application of ranking selection methods used on daily physical activity automatic recognition. An efficient classification method requires a productive and limited feature set, being necessary a selection process since the initial set is quite huge. We have defined a feature selector based on statistical discrimination and robustness criteria, focused on low computational time and resources, defining a real alternative to other selection processes.

For future work, we aim to make a time-based comparison to traditional features selectors [6, 13, 14].

Acknowledgments. This work has been supported by the Spanish CICYT Project TIN2007-60587, Junta de Andalucia Project P07-TIC-02768 and the CENIT project AmIVital, of the "Centro para el Desarrollo Tecnológico Industrial" (CDTI- Spain). We want to express our gratitude to Prof. Stephen S. Intille, Technology Director of the House_n Consortium in the MIT Department of Architecture for the experimental data provided.

References

1. Baek, J., Lee, G., Park, W., Yun, B.J.: Accelerometer Signal Processing for User Activity Detection. In: Negoita, M.G., Howlett, R.J., Jain, L.C. (eds.) KES 2004. LNCS (LNAI), vol. 3215, pp. 1611–3349. Springer, Heidelberg (2004)
2. Bao, L., Intille, S.S.: Activity Recognition from User-Annotated Acceleration Data. In: Ferscha, A., Mattern, F. (eds.) PERVASIVE 2004. LNCS, vol. 3001, pp. 1–17. Springer, Heidelberg (2004)
3. Bonomi, A.G., Goris, A.H.C., Yin, B., Westerterp, K.R.: Detection of Type, Duration, and Intensity of Physical Activity Using an Accelerometer. Medicine & Science in Sports & Exercise 41, 1770–1777 (2009)
4. Crampton, N., Fox, K., Johnston, H., Whitehead, A.: Dance, Dance Evolution: Accelerometer Sensor Networks as Input to Video Games. In: IEEE International Workshop on Haptic, Audio and Visual Environments and Games, pp. 107–112 (2007)
5. Ermes, M., Pärkka, J., Mantyjarvi, J., Korhonen, I.: Detection of daily activities and sports with wearable sensors in controlled and uncontrolled conditions. IEEE Trans. Inf. Technol. Biomed. 12, 20–26 (2008)
6. Kohavi, R., Sommereld, D.: Feature Subset Selection Using the Wrapper Method: Overtting and Dynamic Search Space Topology. In: First International Conference on Knowledge Discovery and Data Mining (1995)
7. Koskimaki, H., Huikari, V., Siirtola, P., Laurinen, P., Roning, J.: Activity recognition using a wrist-worn inertial measurement unit: A case study for industrial assembly lines. In: 17th Mediterranean Conference on Control and Automation, pp. 401–405 (2009)
8. Lee, S.W., Mase, K.: Activity and location recognition using wearable sensors. IEEE Pervasive Computing 1, 24–32 (2002)
9. Mantyjarvi, J., Himberg, J., Seppanen, T.: Recognizing human motion with multiple acceleration sensors. In: Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics, pp. 747–752 (2001)
10. McCue1, M., Hodgins, J., Bargteil, A.: Telerehabilitation in Employment/Community Supports Using Videobased. In: Activity Recognition. RERC on Telereahabilitation (2008)
11. Munguia, E., Intille, S.S., Haskell, W., Larson, K., Wright, J., King, A., Friedman, R.: Real-Time Recognition of Physical Activities and Their Intensities Using Wireless Accelerometers and a Heart Rate Monitor. In: Proceedings of the 2007 11th IEEE International Symposium on Wearable Computers, pp. 1–4 (2007)
12. Ravi, N., Dandekar, N., Mysore, P., Littman, M.L.: Activity Recognition from Accelerometer Data. In: Proceedings of the 17th conference on Innovative applications of artificial intelligence, pp. 1541–1546 (2005)
13. Song, L., Smola, A., Gretton, A., Borgwardt, K.M., Bedo, J.: Supervised feature selection via dependence estimation. In: Proceedings of the 24th international conference on Machine learning, pp. 823–830 (2007)
14. Xu, Z., Jin, R., Ye, J., Lyu, M.R., King, I.: Non-monotonic feature selection. In: Proceedings of the 26th Annual International Conference on Machine Learning, pp. 1145–1152 (2009)