A Novel Feature Extraction Method on Activity Recognition Using Smartphone

Dachuan Wang[1], Li Liu[2], Xianlong Wang[1], and Yonggang Lu[1(✉)]

[1] School of Information Science and Engineering, Lanzhou University,
Gansu 730000, China
ylu@lzu.edu.cn
[2] School of Software Engineering, Chongqing University, Chongqing 400044, China

Abstract. In the last few years, research on human activity recognition using the built-in sensors of smartphones instead of the body-worn sensors has received much attention. Accelerometer is the most commonly used sensor of smartphone for the application. An important step in activity recognition is feature extraction from the raw acceleration data. In this work, a novel feature extraction method which considers both the distribution and the rate of change of the raw acceleration data is proposed. The raw time series liner acceleration data was collected by a smartphone application developed by ourselves. The proposed feature extraction method is compared with a previously proposed statistics-based feature extraction method using two evaluation methods: (a) distance matrix before clustering, (b) ARI and FM-index after clustering using MCODE. Both results show that the newly proposed feature extraction method is more effective for daily activity recognition than the previously proposed method.

Keywords: Activity recognition · Feature extraction · Smartphone · Unsupervised classification

1 Introduction

At present, smartphones are so popular that people carry with them almost anywhere and anytime. Most of the smartphones are equipped with a rich set of embedded sensors, such as accelerometer, GPS sensor, gyroscope, etc. In the last few years, some works using the built-in sensors of smartphones for human activity recognition (HAR) have been proposed. Human activity recognition is one of the important and challenging research areas in ubiquitous computing since it has a wide range of applications including security, healthcare, lifestyle analysis, smart environments, surveillance, etc. Camera-based computer vision systems have been widely used for human activity tracking, but they mostly require infrastructure support, for example, complete camera coverage in the monitoring areas [1]. Alternatively, inertial sensor-based systems for human activity recognition has become an active field of research in the domain of pervasive and mobile computing. Within various sensors, accelerometer is the most commonly used

sensor for recoding body motion signals, because daily activities such as walking, standing, sitting and jogging can be clearly defined by the motion of the body parts. In the work of Bao et al. [2], five biaxial accelerometers are placed in five locations on the user's body to monitor 20 types of activities, trained with 20 users using well-known machine learning classifiers. However, these special sensors are usually not available for the common users. In [3], the acceleration data collected from the embedded triaxial accelerometers of smartphones have been investigated for HAR. As in those different approaches to activity recognition, standard classification algorithms cannot be applied directly on raw time series data. Usually, feature extraction methods have to been used to produce a new data representation (called features) before the classification. Popular features computed from the acceleration data are mean, variance or standard deviation, energy, entropy, correlation between axes or discrete FFT coefficients [4]. In the work by Ravi et al. [5], four features (mean, standard deviation, energy and correlation) are extracted from each axis of a single triaxial accelerometer to recognize eight activities.

In this work, we have proposed a novel feature extraction method which creates a feature vector from the raw time series acceleration data aiming to enhance activity recognition rate especially for the activities ascending stairs and descending stairs. Both the distribution and the rate of the change of the acceleration data are considered in the feature extraction. Molecular Complex Detection (MCODE) clustering method is used to classify these human daily activities including walking, jogging, ascending stairs, descending stairs, sitting and standing. MCODE is an unsupervised clustering method, which has been successfully applied to distinguish race walking from normal walking and running in our previous work [6].

2   Related Work

Many human-activity recognition systems have been proposed such as Camera-based computer vision systems and inertial sensor-based systems. In computer vision-based activity recognition, the common approach is to extract image features from the images or video and to issue a corresponding activity class label [1]. In general, the computer vision-based techniques for human activity tracking often work well in a laboratory or well-controlled environment. However, they require cameras to be placed beforehand at the predetermined points of interest, and they can be influenced by the lighting conditions. Hence, these techniques are not appropriate for highly varied activities that take place in the natural environments. Alternatively, many human activity recognition systems based on the inertial sensor have been developed in recent years. Some of the earliest works based on the inertial sensor focus on the use of multiple accelerometers and possibly other sensors. Bao et al. [2] have used five biaxial accelerometers worn on the user's right hip, wrist, upper arm, ankle and thigh to collect acceleration data from 20 subjects while performing 20 activities. Using several classifiers, they create models to recognize twenty daily activities. It is shown that the decision

tree based classifier shows the best performance for recognizing daily activities, which can produce an overall accuracy rate of 84 %. Furthermore, their results show that placing accelerometers on the subjects thigh is the most effective way for distinguishing the set of 20 activities. Ravi et al. [5] have used a single triaxial accelerometer worn near the pelvic region to distinguish 8 activities: standing, walking, running, ascending stairs, descending stairs, vacuuming, brushing, and situps. They have run several supervised classifiers on data sets in four different settings. It is shown that ascending stairs and descending stairs are hard to be distinguished. In the work of Kwapisz et al. [3], smartphone accelerometers are used to perform activity recognition with data collected from 29 subjects, each carrying a smartphone in their pocket as they perform six activities which are walking, jogging, ascending stairs, descending stairs, sitting and standing, using three supervised classification methods (Decision Trees, Regression, Neural Network). The results show that the activities except ascending stairs and descending stairs can be recognized correctly with a accuracy over 90 %, while the two activities ascending stairs and descending stairs are difficult to be distinguished. All the above studies are using supervised classification methods to perform activity recognition, while in our previous work [6] an unsupervised method named MCODE is used for recognizing race walking using smartphone sensors. The experimental results show that MCODE is effective to distinguish race walking from normal walking and running using smartphone accelerometers.

As mentioned above, in both papers [3,5], it is found that two activities ascending stairs and descending stairs are difficult to be classified. So we have proposed the feature extraction method for improving activity recognition accuracy, especially for the activities ascending stairs and descending stairs.

3 Method

3.1 Data Collection

In order to collect data for our experiment, we have developed a simple Android application that runs on the smartphone. The application permits us to start or stop the data collection and label the activity through a simple graphical user interface. When the data collection is stopped, the data will be saved in a textfile. By setting the application to record the data of linear acceleration sensor type, we can collect linear acceleration data along the x-axis, y-axis and z-axis of the device, not including gravity. Six healthy subjects with ages from 22 to 28 volunteer for the study. Each of them is instructed to carry a Samsung Galaxy SIII smartphone in their front pants leg pocket while performing a specific set of activities for a certain time. These activities are walking, jogging, ascending stairs, descending stairs, sitting and standing. The data collection process is under supervision by one of our team members to ensure the quality of the data.

3.2 Feature Extraction

In the data collection, we have obtained triaxial linear acceleration time series along x-axis, y-axis and z-axis. Figure 1 demonstrates these axes relative to a
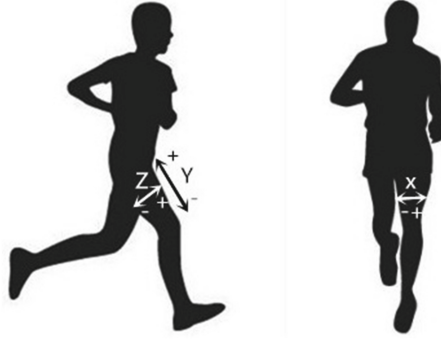
Fig. 1. Axes of motion relative to user

user. In this section, we present a novel method to transform the time series data into a feature vector. At first, the sliding window approach is employed to divide time series data into smaller time segments using a window size of 300 with 150 samples overlapping between consecutive windows, where each segment represents an instance of certain activity. At a sampling frequency of 50 Hz using the smartphone, each instance contains 6-s data readings. We choose the window size of 6 s because it can sufficiently capture cycles in these activities for each segment.

In our feature extraction method (the proposed method), for each segment, a 32-dimensional feature vector is created, denoted by $\boldsymbol{F}$ with $\boldsymbol{F}_i$ indicating the ith dimension value of $\boldsymbol{F}$. Let $\boldsymbol{Y}$ denote the 300 samples along y-axis of each segment, with $\boldsymbol{Y}_j$ indicating the jth dimension value of $\boldsymbol{Y}$ in the unit of m/s^2. Let $\boldsymbol{Z}$ denote the 300 samples along z-axis of each segment, with $\boldsymbol{Z}_j$ indicating the jth dimension value of $\boldsymbol{Z}$ in the unit of m/s^2. The $\boldsymbol{F}$ is calculated as follow:

$$\boldsymbol{F}_i = \sqrt{\sum_{j=1}^{300} \exp(-|\boldsymbol{Y}_j - i + 15|)} \qquad i \in [1, 2, ..., 30]$$

$$\boldsymbol{F}_{31} = \sqrt{\sum_{j=2}^{300} |\boldsymbol{Y}_j - \boldsymbol{Y}_{j-1}|}$$

$$\boldsymbol{F}_{32} = \sqrt{\sum_{j=2}^{300} |\boldsymbol{Z}_j - \boldsymbol{Z}_{j-1}|}$$

The first formula gives 30 features which represent the distribution of the acceleration data around 30 distinct acceleration values from -15 m/s^2 to 14 m/s^2 along y-axis, the second and the third formulas give the 31th and 32th features which represent the rate of change of the acceleration data along y-axis and z-axis respectively.

3.3   MCODE

After obtaining the features, the MCODE-based method was utilized for activity recognition. MCODE is an efficient graph clustering method and is first applied to large protein-protein networks [7]. The input to MCODE is an undirected graph and the output are clusters of instances. In order to utilize the MCODE clustering algorithm for activity recognition, all the feature vectors of each subjects data need to be mapped into an undirected graph, denoted by G. The graph G is a complete graph and each vertex in the graph represents a feature vector, namely, an instance of certain activity. The set of all vertices and all edges in G are denoted as V and E, respectively. The length of each edge in E refers to the Euclidean distance between the two corresponding vertices and the edge weight is the reciprocal of the length of the edge. For using the MCODE clustering algorithm effectively, some edges in the complete graph G need to be deleted, which is accomplished by defining a threshold for the edge weight. The value of the threshold is the averaged weight of E multiplied by a parameter named EWP. If the weight of an edge is less than the threshold, the edge will be deleted from the complete graph.

MCODE algorithm operates in three stages, vertex weighting, cluster finding and optionally post-processing. Post-processing is used to assign vertices to multiple clusters and delete the clusters with a small number of vertices [7]. In our study, every vertex belongs to one and only one cluster, so the stage of post-processing is not required. The first stage of MCODE, vertex weighting, weights all vertices based on their local network density using the highest k-core of the vertex neighborhood. Density of a graph, $G = (V, E)$, with number of vertices $|V|$ and number of edges $|E|$, is defined as $|E|$ divided by the number of edges of a complete graph composed of V. Thus, density of G, $DG = |E|/(|V|(|V|-1)/2)$. A k-core is a graph of minimal degree k (graph G, for all v in G, $deg(v) \geq k$). For any vertex $v \in V$, the highest k-core called M is found from the local network N_v that is composed of v and its neighbors. Finally, the weight of v is the product of Ms k-core and Ms density. The second stage, cluster finding, takes as input the vertex weighted graph, seeds a cluster with the highest weighted vertex and set the vertex with the highest weight as seed vertex. Then for each neighbor v of the seed vertex, if the weight of v is greater than a given threshold, which is the weight of the seed vertex multiplied by a parameter named VWP, the v will be classified into the cluster which the seed belongs to and set as new seed. Then the neighbours of the new seed will be recursively checked in the same manner. A vertex is not checked more than once and this process stops once no more vertices can be added to the cluster based on the given threshold and is repeated for the next highest unseen weighted vertex in the graph.

4   Results and Discussion

In our previously proposed method [6], statistic-based features are extracted, which includes mean, standard deviation, variance, skewness, kurtosis, signal magnitude area and correlation between axis pair. In this work, we compared the

previous method with the newly proposed feature extraction method. Adjusted Rand Index (ARI) [8] and Fowlkes-Mallows Index (FM-index) [9] were used to evaluate the experimental results. The ARI and FM-index are both used to measure the similarity between the clustering result and the benchmark. The maximum value of the ARI and FM-index are both 1, which means that the clustering result are exactly the same as the benchmark, while the minimum value of them are both 0, which means the clustering result is a random result. For both the ARI and FM-index, a larger value means a better similarity between the clustering result and the benchmark.

4.1   Parameter Selection Using ARI

As mentioned above, two parameters, EWP and VWP, need to be provided for applying the MCODE method for activity recognition. In order to find the best parameters for both methods, we have tested various parameter combinations of EWP and VWP on the dataset of 6 subjects. Obviously, the largest sum of ARI values of 6 subjects indicates the best parameter setting. Thus, with each combination of EWP and VWP, for each subjects data, we have computed the ARI value of clustering result by applying the proposed method and the previous method, respectively. Figure 1(a) shows the sum of ARI values of different combinations of EWP and VWP produced by the proposed method, while Fig. 1(b) shows the sum of ARI values produced by the previous method. As shown in Fig. 2, when EWP = 1.6 and VWP = 0.1, the sum of ARI values reaches the largest for the proposed method, whereas the parameter setting of EWP = 1.8 and VWP = 0.1 leads to the largest sum of ARI values for the previous method. In the experiments, it is found that the best parameter settings given by FM-index are exactly the same as the results listed above. Therefore, we have selected EWP = 1.6 and VWP = 0.1 for the proposed method, while EWP = 1.8 and VWP = 0.1 are used for the previous method.

4.2   Evaluation

Evaluation Using Distance Matrix. A good feature extraction method for activity recognition ought to make sure that the exacted features of the same activity are similar while the features of different activities are dissimilar. In other words, the more similar the exacted features of the same activity, and the more dissimilar the features of the different activities, the better the feature extraction method is. So, the pairwise Euclidean distance matrix of the feature vectors can be used to measure the similarities between features and thus can be used to evaluate the feature extraction method. Here the Euclidean distance matrixes are used to compare the proposed feature extraction method with previous method. For each subjects data, we have constructed two Euclidean distance matrixes produced using the proposed method and previous method respectively. Each matrix is normalized by dividing the mean value of the matrix. All the matrixes are shown as color map in Fig. 3 through Fig. 8. In all the figures, a pixel represents the distance between the two corresponding instances of activities so that
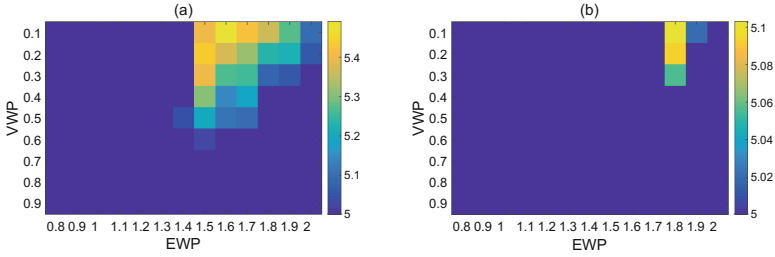
Fig. 2. Sum of ARI values for different combinations of EWP and VWP. (a) Proposed Method; (b) Previous Method
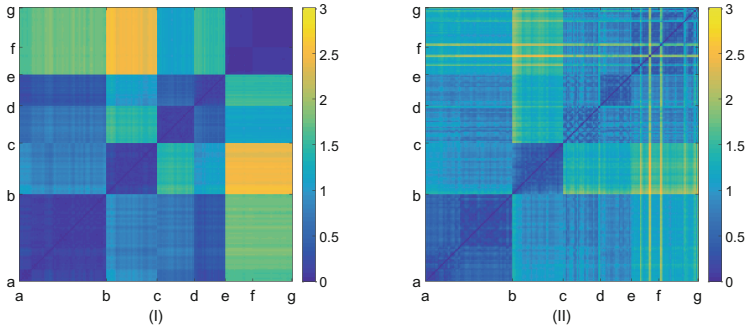

Fig. 3. Distance matrix of activities (walking: a–b, jogging: b–c, ascending stairs: c–d, descending stairs: d–e, sitting: e–f, standing: f–g) of *Subject 1* with (I) Proposed Method; (II) Previous Method
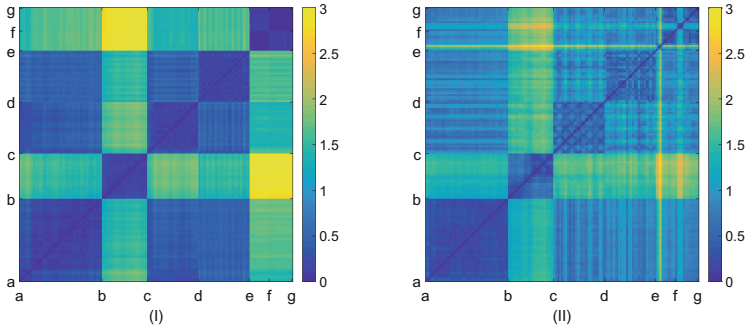

Fig. 4. Distance matrix of activities (walking: a–b, jogging: b–c, ascending stairs: c–d, descending stairs: d–e, sitting: e–f, standing: f–g) of *Subject 2* with (I) Proposed Method; (II) Previous Method

the area of the squares of each activity is determined by the square of the number of instances of the activity.

Obviously, the proposed feature extraction method outperforms previous method as shown in each figure. It is clearly shown that the similarities within
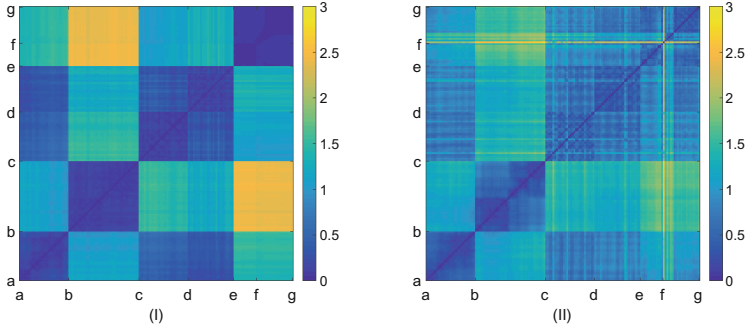
Fig. 5. Distance matrix of activities (walking: a–b, jogging: b–c, ascending stairs: c–d, descending stairs: d–e, sitting: e–f, standing: f–g) of *Subject 3* with (I) Proposed Method; (II) Previous Method
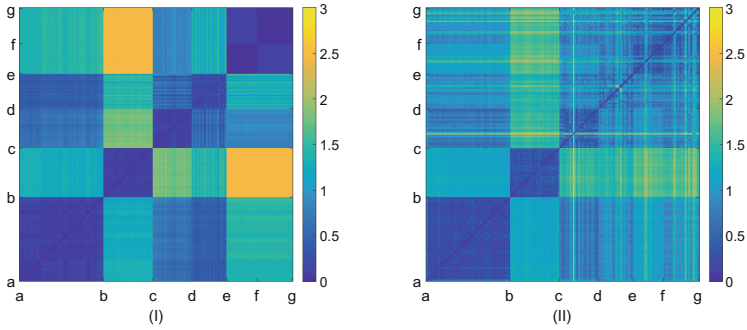

Fig. 6. Distance matrix of activities (walking: a–b, jogging: b–c, ascending stairs: c–d, descending stairs: d–e, sitting: e–f, standing: f–g) of *Subject 4* with (I) Proposed Method; (II) Previous Method

the same activity for the proposed method is apparently higher than the previous method in all the figures. For the proposed method, all the figures show that the activities including walking, jogging, ascending stairs and descending stairs are all distinct from each others, while for the previous method, ascending stairs and descending stairs cannot be easily distinguished in some cases, such as in Figs. 5, 6 and 8. These results show that the proposed feature extraction method is more effective for distinguishing between ascending stairs and descending stairs. It is also found that the sitting and standing are not distinct from each other in the results of both methods. It is well known that both sitting and standing are static, namely, the liner acceleration of them are all near 0, and our approach only uses the single liner acceleration readings of activities, which may explain the failure for distinguishing the two activities, which is also the case for the previous method. For both of the two methods, it can be seen from Fig. 4 that the similarities among the walking, ascending stairs and descending stairs are relatively high, which is also mentioned in the work [10].
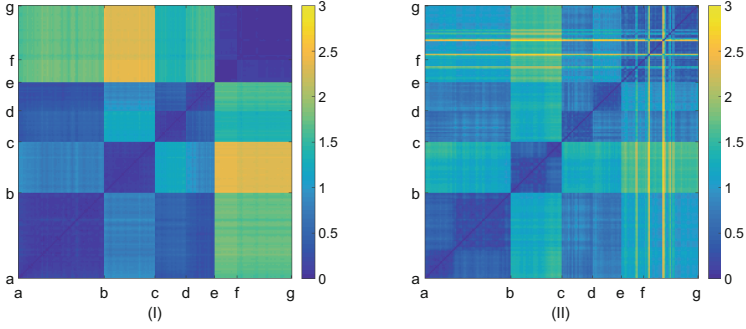
Fig. 7. Distance matrix of activities (walking: a–b, jogging: b–c, ascending stairs: c–d, descending stairs: d–e, sitting: e–f, standing: f–g) of *Subject 5* with (I) Proposed Method; (II) Previous Method
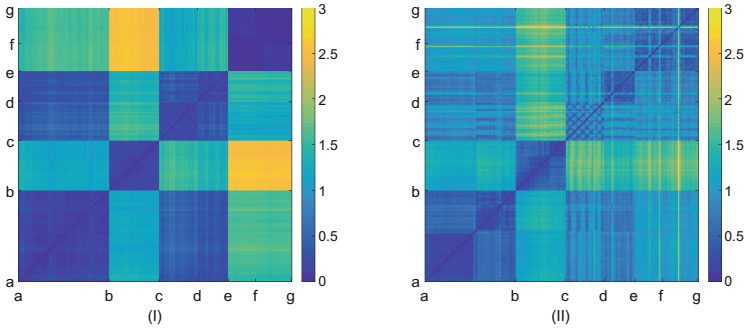

Fig. 8. Distance matrix of activities (walking: a–b, jogging: b–c, ascending stairs: c–d, descending stairs: d–e, sitting: e–f, standing: f–g) of *Subject 6* with (I) Proposed Method; (II) Previous Method

Table 1. Comparison of the proposed feature extraction method with the previous method using ARI and FM-index

		Subject1	Subject2	Subject3	Subject4	Subject5	Subject6
ARI	Proposed Method	0.9119	0.9363	0.9264	0.9487	0.8859	0.8818
	Previous Method	0.8841	0.9135	0.7443	0.8313	0.8700	0.8597
FM-index	Proposed Method	0.9326	0.9490	0.9420	0.9589	0.9120	0.9068
	Previous Method	0.9065	0.9308	0.8044	0.8630	0.8963	0.8874

Evaluation Using ARI and FM-index. The best parameters selected for both the proposed method and the previous method in Sect. 4.1 are used in the evaluation, which are EWP = 1.6 and VWP = 0.1 for the proposed method, and EWP = 1.8 and VWP = 0.1 for the previous method. The ARI and FM-index produced by the two methods for 6 subjects are listed in Table 1. It can be seen that for each subject, both the ARI value and the FM-index value produced

by the proposed method are higher than these of the previous method. These results show that the proposed method is more effective for the daily activity recognition than the previous method (Fig. 7).

5   Conclusion

In this paper a novel feature extraction method for activity recognition based on the distribution and the rate of change of the linear acceleration data is proposed. Activities except for sitting and standing can be distinguished from each other by the proposed method. Compared to a previously proposed statistics-based feature extraction method, the newly proposed method is shown to be more effective for daily activity recognition, especially for distinguishing between ascending stairs and descending stairs. In the future work, we will involve more phone-based sensors than linear acceleration senor because the liner acceleration are not enough to separate the statical activities such as sitting and standing of which the liner acceleration are both close to zero.

References

1. Poppe, R.: A survey on vision-based human action recognition. Image Vis. Comput. **28**, 976–990 (2010)
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). doi:10.1007/978-3-540-24646-6_1
3. Kwapisz, J., Weiss, G., Moore, S.: Activity recognition using cell phone accelerometers. ACM SigKDD Explor. Newslett. **12**(2), 74–82 (2011)
4. Huynh, T., Schiele, B.: Analyzing features for activity recognition. In: Proceedings Conference on Smart Objects and Ambient Intelligence, pp. 159–163. ACM, New York (2005)
5. Ravi, N., Dandekar, N., Mysore, P., Littman, M.: Activity recognition from accelerometer data. Am. Assoc. Artif. Intell. **05**, 1541–1546 (2005)
6. Wei, Y., Liu, L., Zhong, J., Lu, Y., Sun, L.: Unsupervised race walking recognition using smartphone accelerometers. In: Zhang, S., Wirsing, M., Zhang, Z. (eds.) KSEM 2015. LNCS (LNAI), vol. 9403, pp. 691–702. Springer, Heidelberg (2015). doi:10.1007/978-3-319-25159-2_63
7. Bader, G., Hogue, C.: An automated method for finding molecular complexes in large protein interaction networks. BMC Bioinf. **4**(1), 2 (2003)
8. Lawrence, H., Arabic, P.: Comparing partitions. J. Classif. **2**(1), 193–218 (1985)
9. Fowlkes, E., Mallows, C.: A method for comparing two hierarchical clusterings. J. Am. Stat. Assoc. **78**(383), 553–569 (1983)
10. He, Y., Li, Y.: Physical activity recognition utilizing the built-in kinematic sensors of a smartphone. Int. J. Distrib. Sens. Netw. **2013**, 10 pages (2013). doi:10.1155/2013/481580. Article ID 481580