



Contents lists available at SciVerse ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa



Online motion recognition using an accelerometer in a mobile device

D. Fuentes^a, L. Gonzalez-Abril^{b,*}, C. Angulo^c, J.A. Ortega^a

^aComputer Languages and Systems Dept., University of Seville, 41012 Seville, Spain

^bApplied Economics I Dept., University of Seville, 41018 Seville, Spain

^cTechnical Research Centre for Dependency Care and Autonomous Living – CETpD., Universitat Politècnica de Catalunya, 08034 Vilanova i la Geltru, Spain

ARTICLE INFO

Keywords:
Features extraction
Pattern recognition
SVM

ABSTRACT

This paper introduces a new method to implement a motion recognition process using a mobile phone fitted with an accelerometer. The data collected from the accelerometer are interpreted by means of a statistical study and machine learning algorithms in order to obtain a classification function. Then, that function is implemented in a mobile phone and online experiments are carried out. Experimental results show that this approach can be used to effectively recognize different human activities with a high-level accuracy.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

Motion recognition is a discipline that has been around lately in the human–computer interaction research community (Nayak, Sarkar, & Loeding, 2009; Parera, Angulo, & Rodriguez-Moliner, 2009). Motion detection is a similarly hard problem as speech or gesture recognition since these are complex acts in human beings. Nevertheless, this field has a wide variety of applications. For example, a continued study of this information can help a doctor to establish a correct diagnosis or a rehabilitation plan for persons with mobility problems.

Many mobile phones can be found nowadays fitted with devices such as accelerometers, gyroscopes, cameras, etc., enabling us to obtain data from their users such as their movements. Furthermore, these mobile phones have a really high processing capacity to execute complex programs. This feature allows us the implementation of reliable methods to recognize movements using the data from the accelerometer. However, that recognition process implies a common problem: human movements are very complex because many actions can take place both sequentially and simultaneously. Due to that, there are many different combinations of sequential and simultaneous human movements so it is almost impossible to model them all explicitly. Using an accelerometer with a person, the data collected can be different depending for example on the age of the person who stands up, i.e., a young boy or an elder person.

Another similar situation could be considered when two persons with the same age do a lateral movement and one of them has one of his/her legs injured. To begin with, the time in which these persons would do the movement would not be the same.

Secondly, if these persons have a device fitted with accelerometers the data captured would be different. Hence, a recognition method has to be designed in order to recognize movements in persons of different sex, age or physical condition. Another problem is where to place the recognition device, since the data retrieved from the accelerometer can be different if the same person has the device in a different pocket of his trousers or even in the same pocket.

From the medical point of view, many works have been developed using the data retrieved from the accelerometers' sensors to study chronic diseases, strokes or rehabilitation processes. Accelerometers have a high potential use in monitoring patients undergoing rehabilitation processes because the information provided together with a clinical assessment can help shortening the duration of the rehabilitation plan by applying the appropriate therapeutic intervention earlier. Clinicians and biomedical engineers are joining forces to make this technology part of the routine in clinical practice (Bobick & Davis, 2001; Sminchisescu, Kanaujia, Li, & Metaxas, 2006). Hence, the processing of data from the accelerometer allows determining the rehabilitation activities a patient would require. This information, for example, could on the one hand help a doctor to determine if a patient is doing his/her exercises correctly or, on the other hand, it could be taken as a sign that helped to correct the rehabilitation plan.

The acceleration signals recorded via the accelerometers have been used in many works to classify daily life activities (such as sitting, standing and even walking) (Li & Ge, 2009; Parera et al., 2009). Most notable accelerometer studies discriminate one movement from other activities, e.g., fall detection (Zhang, Wang, Liu, & Hou, 2006). Some methods use other technologies rather than accelerometry such as RFID (Alvarez, Perez, Angulo, & Ortega, 2007; Hein & Kirste, 2007) or video recording (Matta & Dugelay, 2009). Likewise, a wide variety of classification methods are used: KFD algorithms (Zhang et al., 2006), Bayesian and neural networks

* Corresponding author.

E-mail address: luisgon@us.es (L. Gonzalez-Abril).

(Bu, Okamoto, & Tsuji, 2009), support vector machines (Cao, Masoud, Boley, & Papanikolopoulos, 2009; Parera et al., 2009) or decision trees (Rogez, Rihan, & Ramalingam, 2008). All these works describe online recognition, consisting of collecting data and then recognizing the movements. However, in this paper, we describe a new solution and implement an online method to recognize human movements in real time using a phone with an accelerometer.

The rest of this paper is structured as follows: Section 2 describes the data captured and the activities studied. The feature selection strategy based on discriminate power for generated features is presented in Section 3. Data interpretation and processing is described in Section 4. Experimentation is carried out in Section 5 and the last section contains the conclusions drawn.

2. Information collection

In this section, the relation between the device used and the way to obtain data from the accelerometer are explained. The activities to be classified are described next.

2.1. Data collection

An accelerometer measures acceleration (measured in m/s^2) caused either by a motion or by the gravity. Acceleration near the surface of the earth is around $9.8 m/s^2$, which is used as unit of measure and is denoted as G , that is, $1 G = 9.8 m/s^2$. The data retrieved from an accelerometer can provide the necessary data to study the behaviour of a person, e.g., the patient movements in his rehabilitation exercises. In this paper, the input signals are obtained from a triaxial accelerometer of the sensor device located on the user's chest sampled at 100 Hz. This position has been chosen because, despite the asymmetry of the human body, usually in both men and women the centre of mass (a pivot point around which the system can revolve (Wikia Education, 2010)) is around the chest. The data from the triaxial accelerometer is represented in 3-column values ($X; Y; Z$) measured in G . The orientation of the three axes is variable and depends on the device; hence it is an important step before doing any test to establish a common axis orientation in order to be able to compare it later with other motion recognition methods or with the same method implemented in different devices. Data from the accelerometer is stored into text files for the next step and then it is normalized (that is, the mean value is zero and the standard deviation value is one) in order to avoid problems with outliers.

2.2. Activities

The final objective of this work is to identify a set of four daily activities performed by the user while he/she is wearing the sensor device. The activities to be classified are (the orientation of XYZ axes are shown in):

- Sitting down: this movement ranges from 1 to 2.5 s.
- Standing up: this action also ranges from 1 to 2.5 s. It is similar to an inverse action of sitting down.
- Walking: this activity includes taking various steps.
- Stop: the stop motion is considered as the steady standing and sitting action. In both cases the mobile phone stays in the same position, and hence the data obtained is very similar and the movements are considered to be the same.

Fig. 1 shows an example of the Z-axis readings of an accelerometer for each activity. It is worth noting that the duration of the movement depends on the person who carries out the movement.

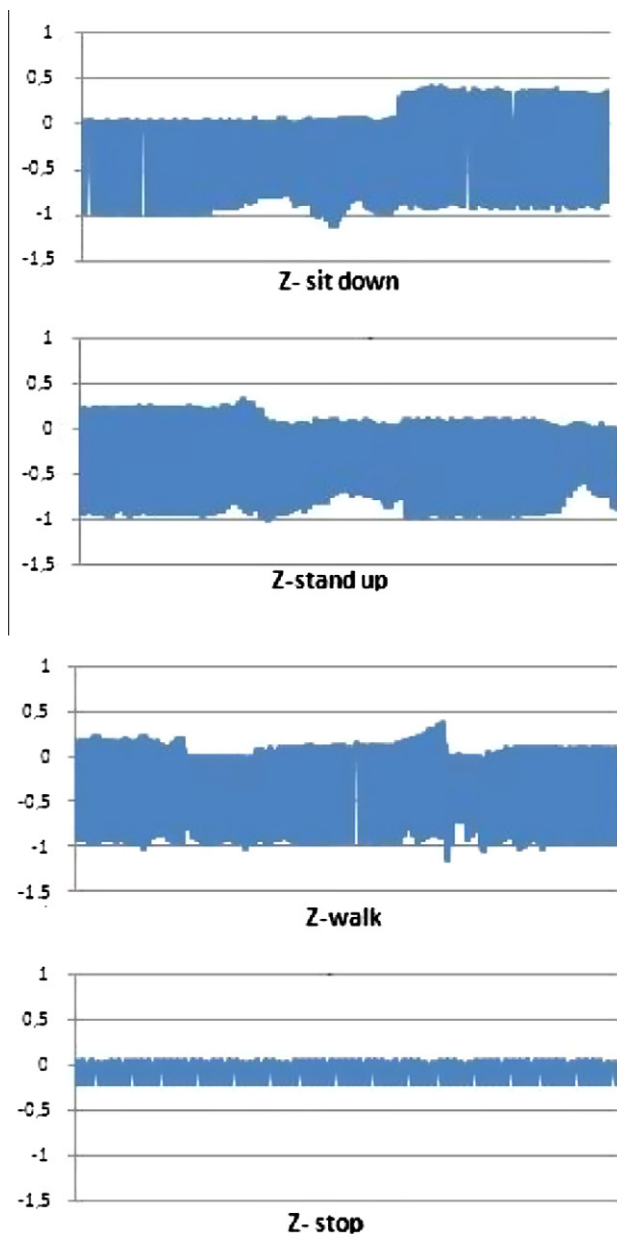


Fig. 1. Z-axis readings for different movements.

For example, a young man can take less than 1 s to stand up while in an elder person the time can be equal to or longer than 1.5–2 s.

3. Features extraction

Once the data is collected, more than 100 related features were considered such as angle calculations, the acceleration module, increments and averages. A deep statistical analysis was carried out in Parera et al. (2009) and previous data analysis tools were used (e.g., boxplot, the Fourier transform, etc.) to choose the next nine features.

- The standard deviation and the range of the orientation θ angle. This angle based on the earth gravity allows computing the sensor device orientation.

$$\text{std}(\theta), \text{range}(\theta) | \theta = \arctan \left(X, \sqrt{Y^2 + Z^2} \right). \quad (1)$$

Q1

3

- The standard deviation and the minimum value of the forward acceleration AF . This acceleration uses the θ angle to compute the accelerations in the earth fixed reference frame by applying the rotation matrix of the X axis.

$$std(AF), range(AF)|AF = \cos(\theta) * X + \sin(\theta) * Z. \quad (2)$$

- The standard deviation of the vertical acceleration AV . The vertical acceleration is similar to the forward acceleration and it uses the θ and ψ orientation angles.

$$std(AV)|AV = \cos(\psi) * X + \sin(\theta) * X + \sin(\psi) * Y + \cos(\psi) * Z, \psi = \arctan(Y, Z). \quad (3)$$

- The standard deviation and the minimum values of the Y values of the window.

$$std(Y), range(Y). \quad (4)$$

- The standard deviation and the minimum values of the Z values of the window.

$$std(Z), range(Z). \quad (5)$$

The orientation of the XYZ axis are shown in Fig. 4. The features selected depend on the activities studied. In this case, none of these activities implies a lateral movement. In the accelerometer, lateral movements are represented by the X axis; hence the features which include Y or Z axes are better classifiers.

4. Data labeling and processing

As aforementioned, the values from accelerometer are recorded into a 3-column format file. Once the movement is finished, the data file is divided into different windows of 100-xyz values. The nine statistical features are computed for each window and the results are recorded in another file. Then, an automatic procedure is carried out to decide which ones will be useful in the classification machine learning process. In activities involving stopping and walking, all entries (we consider an entry a set of nine statistical results) are considered useful for machine learning training and they determine the learning set. The situation for the stop movement (same as for walking) is shown in Fig. 2.

However, the cases in the standing up and sitting down activities are different. Before a person sits down (or stands up) he remains steady in a standing up position (or in a seating position) and, after that person has finished that motion, he sits down (or remains steady in a standing up position). Hence, the beginning and ending parts of the movements are more similar to the stop movement, and consequently, the activity label for these times may not correspond to the actual times of the activity performed. Therefore, the entries which are located in the centre of the movement are

considered for the learning set. These entries are labelled with the movement and they will be useful to recognize standing up and sitting down movements. Nevertheless, the other entries computed with the data gathered since the beginning of the motion until its end are ruled out because they can generate errors due to their similarity with the stop activity. Fig. 3 describes this procedure.

Once the training period ends and the features are computed, an estimation of the patient behaviour is stored in the device. After that, the aim is to determine the activity associated with a new entry with nine values from the statistical features using a supervised learning method. In this paper tests have been carried out specifically using a support vector machine (SVM) (Gonzalez, Angulo, Velasco, & Catala, 2006; Vapnik, 1998) which is based on $1 - \nu - 1$ SVM with different normalization outputs (Gonzalez-Abril, Angulo, Velasco, & Ortega, 2010).

Using this configuration, the SVM learns the behaviour of a patient derived from the data of the accelerometer through hundreds of labelled instances which are formed by nine variables and a label. In a first step, the SVM is trained by multiple labelled entries. After the execution, the result obtained from training a SVM is a trained classification function which will help to classify new examples without label. The program computes automatically the classification function with the training set and tests it with the testing set. The best cross-validation mean rates (the accuracy) (C, σ^2) and the necessary parameters to implement directly the classification function in the mobile phone are reported in the next section.

5. Experimentation

In this section, we present the results from experimentation. In order to increase the robustness of the system, the effectiveness of the method has been studied on ten men and women between 20 and 55 years old with different mobility limitations. For this task, the Openmoko mobile phone Neo FreeRunner (Openmoko Wiki, 2010) was used to implement and test our system. Openmoko is a Linux distribution designed for open mobile computing platforms, such as, but not limited to, mobile phones. It is not tied to any particular mobile phone providing developers with the capability to easily create and deploy applications. Openmoko is currently selling the Neo FreeRunner model, in which there are two three-axes accelerometer sensors and the information from them is exported through two different input event based file mappings. A program has been designed to take the retrieval signals from one of the accelerometers whose orientation is showed in Fig. 4.

This experimentation section has been split in offline and online recognition tests. The main difference between them is based on the decision to implement the classification function into the mobile device. In the offline experimentation, the phone is only needed to obtain the training and testing sets.



Fig. 2. Features calculation sample for stopping (or walking) activity.

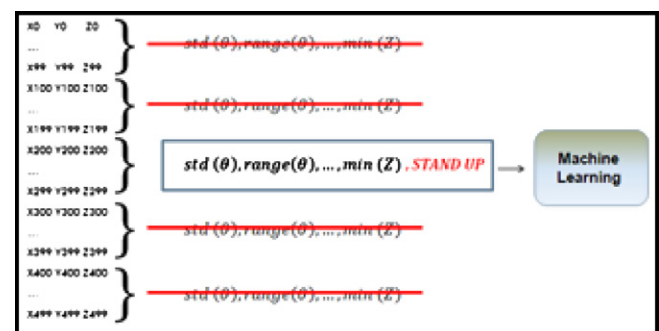


Fig. 3. Features calculation sample for standing up (or sitting down) activity.



Fig. 4. Orientation of the axes in the Openmoko Neo FreeRunner accelerometer used.

a two-dimensional grid with the following values: $C = \{2^{10}, 2^9, \dots, 2^{-2}\}$ and $\sigma^2 = \{2^4, 2^3, \dots, 2^{-4}\}$. The criteria used to estimate the generalized accuracy is the 10-fold cross-validation on the whole set of training data and this procedure is repeated 50 times in order to ensure a good statistical behaviour. The optimization algorithm used is the exact quadratic program-solver provided by Matlab. Results have also been normalized in order to prevent problems with outliers. Data from the test has been normalized.

The best cross-validation mean rate among the pairs (C, σ^2) and its standard deviation are reported in Table 1.

The best accuracy rate is 93.4% which is attained with the values $C = 2^5$ and $\sigma^2 = 2$, and a standard deviation equal to 0.12. In order to test whether there were statistically significant differences between the accuracy rates, hypothesis tests were carried out:

Let $X_i = 0$ be assigned when the machine provide an error for the pattern x_i and $X_i = 1$ otherwise. Then X_i follows a Bernoulli p.d.f., $X_i \sim B(p_i)$, where p_i is the probability that the machine assigns a correct class for a new input. By testing $H_0 : p_1 = p_2$ versus $H_1 : p_1 \neq p_2$, the statistics is $Z = \frac{\sqrt{n}(\bar{X}_1 - \bar{X}_2)}{\sqrt{\bar{X}_1(1-\bar{X}_1) + \bar{X}_2(1-\bar{X}_2)}} \sim N(0, 1)$ and \bar{X} is the mean sample of size n .

Hence, no statistically significant difference between accuracy rates was found.

Therefore, according to the empirical experimentation carried out in this work, some analysis can be completed: (i) the accuracy rate can be improved using different values of C and σ^2 , although no statistically significant differences between accuracy rates were found. (ii) The standard deviation of the mean accuracy rate is very similar for each considered classification problem, which is a logical result since the optimization problem for all the iterations is the same. (iii) The standard deviation depends on the mean accuracy rate more than on any other characteristic of the data set, which is a natural theoretical result. (iv) It can be observed that the difference in performance between parameters is small.

5.2. Online recognition

Once the SVM classification function is obtained and the best C and σ^2 values are computed in the offline recognition, the function is implemented into the mobile phone with the objective of labelling new entries in real time. To demonstrate the effectiveness of the online method, the test group repeated each activity 10 times. In the phone, the data is collected, the features are computed to form a new entry and, using the classification function, the label is assigned, then it is displayed on the screen of the phone and finally it is stored in a log file. Tests results are listed in Table 2.

In all these experiments the Openmoko accelerometers were noticed to be very sensitive to light vibrations, and the stop motion

5.1. Offline recognition

For the offline experimentation, the dataset is obtained in first place. The test group performed each activity 10 times and 1300 labelled entries were computed in total. Using this dataset, a Matlab program was used for the offline recognition. This application implements the SVM where C and σ^2 values are configurable parameters and the size of labelled-training and the unlabelled-testing sets are variable. For each C and σ^2 values, the SVM is trained with the training set in order to obtain a function able to recognize new entries. Then, the testing set is used to recognize movements, each entry is labelled by the function and the accuracy is computed.

Due to its spread use, the performance for the multi-classification approach, in the form of the accuracy rate has been evaluated on models using the Gaussian kernel, $k(x_i, x_j) = \exp^{-\|x_i - x_j\|^2 / 2\sigma^2}$, and therefore two hyper-parameters must be set: the regularization term C and the width of the kernel σ^2 . This space is explored on

Table 1 Accuracy (%) ± Standard deviation for different values of the parameters C and σ^2 .

C	σ^2 2 ⁴	2 ³	2 ²	2 ¹	2 ⁰	2 ⁻¹	2 ⁻²	2 ⁻³	2 ⁻⁴
2 ¹⁰	92.4 ± 0.14	92.6 ± 0.13	92.5.1 ± 0.12	92.6 ± 0.14	92.4 ± 0.12	92.3 ± 0.13	92.5 ± 0.11	92.5 ± 0.15	92.4 ± 0.12
2 ⁹	92.5 ± 0.12	92.4 ± 0.13	92.4 ± 0.12	92.3 ± 0.12	92.4 ± 0.14	92.6 ± 0.13	92.3 ± 0.15	92.5 ± 0.14	92.3 ± 0.15
2 ⁸	92.7 ± 0.12	92.6 ± 0.12	92.5 ± 0.14	92.5 ± 0.16	92.6 ± 0.12	92.4 ± 0.13	92.4 ± 0.13	92.4 ± 0.13	92.4 ± 0.12
2 ⁷	92.8 ± 0.14	92.8 ± 0.14	92.7 ± 0.16	93.0 ± 0.15	93.0 ± 0.13	92.8 ± 0.14	92.7 ± 0.12	92.6 ± 0.12	92.8 ± 0.14
2 ⁶	93.1 ± 0.13	92.9 ± 0.13	92.7 ± 0.12	92.9 ± 0.13	93.1 ± 0.12	92.8 ± 0.13	92.7 ± 0.13	92.9 ± 0.14	93.1 ± 0.12
2 ⁵	92.8 ± 0.12	93.0 ± 0.13	92.9 ± 0.14	93.4 ± 0.12	93.1 ± 0.15	93.1 ± 0.12	93.2 ± 0.14	93.1 ± 0.15	93.1 ± 0.12
2 ⁴	(93.0%, 0.13)	93.2 ± 0.13	93.1 ± 0.16	92.9 ± 0.14	93.1 ± 0.13	93.2 ± 0.11	93.1 ± 0.14	93.3 ± 0.15	93.3 ± 0.12
2 ³	92.3 ± 0.14	93.0 ± 0.15	93.2 ± 0.16	93.2 ± 0.14	93.1 ± 0.15	92.8 ± 0.14	93.1 ± 0.14	93.0 ± 0.14	92.9 ± 0.13
2 ²	92.7 ± 0.13	92.8 ± 0.12	92.9 ± 0.13	92.9 ± 0.12	93.1 ± 0.12	93.0 ± 0.14	93.1 ± 0.12	92.7 ± 0.12	92.9 ± 0.14
2 ¹	92.8 ± 0.13	92.9 ± 0.14	92.7 ± 0.13	92.8 ± 0.15	92.9 ± 0.14	92.8 ± 0.13	92.6 ± 0.12	92.6 ± 0.15	92.5 ± 0.13
2 ⁰	92.5 ± 0.15	92.6 ± 0.13	92.8 ± 0.15	92.5 ± 0.12	92.6 ± 0.12	92.9 ± 0.13	92.5 ± 0.15	92.8 ± 0.13	92.6 ± 0.15
2 ⁻¹	92.6 ± 0.13	92.7 ± 0.13	92.8 ± 0.13	92.6 ± 0.14	92.8 ± 0.13	92.6 ± 0.12	92.6 ± 0.14	92.7 ± 0.14	92.5 ± 0.14
2 ⁻²	92.5 ± 0.13	92.7 ± 0.14	92.4 ± 0.14	92.6 ± 0.14	92.4 ± 0.14	92.6 ± 0.13	92.8 ± 0.14	92.4 ± 0.13	92.6 ± 0.14

Q1

5

Table 2
Confusion matrix.

Activity	Stopping	Walking	Standing-up	Sitting-down
Stopping	100	0	0	0
Walking	0	95	2.6	2.4
Standing-up	0	8.3	91.7	0
Sitting-down	0	7.8	0	92.2

was displayed on the screen of the device with 100% accuracy. In relation to the walking movement, each person carried out 10 tests, consisting of taking five steps in a straight line, obtaining 95% accuracy, that is, the mobile phone detects this motion from the moment the person starts moving until he/she stops. In the standing-up and sitting-down activities, each person carried out 10 repetitions using the same chair but taking different time to complete the activity. The total accuracy for these activities was 91.7% and 92.2% respectively. The best accuracy was achieved for the stop and walking activities; during the exercise sessions we found that activities with limited or repeated actions like stopping or walking can be better distinguished from other activities which imply different movements such as standing up or sitting down thanks to the good sensitivity and specificity. Moreover, it was observed that the time for doing activities like walking was usually the same whereas for standing-up or sitting-down it depended on the age or physical condition of the person who performed the activity. Table 2 shows that most errors arose from confusing standing up or sitting down motions with walking and these mistakes can be related to the speed of the movement and the body inclination.

6. Conclusions

In this paper, we have presented a new methodology to recognize motions either offline or online based on data from an accelerometer in a mobile device. Using this method, health personnel can obtain new information about the patient to decide a diagnosis. Both online and offline designed algorithms showed an overall accuracy of 93% when recognizing different activities.

7. Uncited reference

Nguyen, Shin, Shin, and Kim (2009).

Acknowledgment

This research is partially supported by the MCI I + D project ARTEMISA (TIN2009–14378-CO2–01).

References

- Alvarez, J. A., Perez, C., Angulo, C., & Ortega, J. A. (2007). Combining smart tags and body fixed sensors for disabled people assistance. *Lecture Notes in Computer Science*, 4693, 10–17.
- Bobick, A. F., & Davis, J. W. (2001). The recognition of human movement using temporal templates. *The Recognition of Human Movement Using Temporal Templates*, 23(3), 257–267.
- Bu, N., Okamoto, M., & Tsuji, T. (2009). A hybrid motion classification approach for EMG-based human-robot interfaces using Bayesian and neural networks. *IEEE Transactions on Robotics*, 25(3), 502–511.
- Cao, D., Masoud, O. T., Boley, D., & Papanikolopoulos, N. (2009). Human motion recognition using support vector machines. *Computer Vision and Image Understanding*, 113(10), 1064–1075.
- Gonzalez-Abril, L., Angulo, C., Velasco, F., & Ortega, J. A. (2010). Analysis of output normalization in multiclass SVMs. In revision.
- Gonzalez, L., Angulo, C., Velasco, F., & Catala, A. (2006). Dual unification of bi-class support vector machine formulations. In *Pattern Recognition* Vol. 39, No. 7 (pp. 1325–1332).
- Hein, A., & Kirste, T. (2007). A hybrid approach for recognizing ADLs and care activities using inertial sensors and RFID. In *Proceedings of third intl. conf. intelligent sensors, sensor networks and information processing* (pp. 1–7).
- Li, T., & Ge, M. (2009). Human motion recognition using ultra-wideband radar and cameras on mobile robot. In *Transactions of Tianjin University* (Vol. 15, pp. 381–387).
- Matta, F., & Dugelay, J. L. (2009). Person recognition using facial video information: A state of the art. *Journal of visual languages and computing*, 20(3), 180–187.
- Nayak, S., Sarkar, S., & Loeding, B. (2009). Distribution-based dimensionality reduction applied to articulated motion recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(5).
- Nguyen, Q. C., Shin, D., Shin, D., & Kim, J. (2009). Real-time human tracker based on location and motion recognition of user for smart home. In *Proceedings of the third international conference on multimedia and ubiquitous engineering* (pp. 243–250).
- Openmoko Wiki, Openmoko Main Page, (2010). <<http://wiki.openmoko.org/wiki/MainPage>>.
- Parera, J., Angulo, C., & Rodriguez-Molinero, A. (2009). User daily activity classification from accelerometry using feature selection and SVM. *Lecture Notes in Computer Science*, 5517, 1137–1144.
- Rogez, G., Rihan, J., & Ramalingam, S. (2008). Randomized trees for human pose detection. In *IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1–8).
- Sminchisescu, C., Kanaujia, A., Li, Z., & Metaxas, D. (2006). Conditional models for contextual human motion recognition. In *Computer Vision and Image Understanding*.
- Vapnik, V. N. (1998). *Statistical learning theory*. Editorial John Wiley & Sons.
- Wikia Education, Center of mass, (2010). <<http://schools.wikia.com/wiki/CenterofMass>>.
- Zhang, T., Wang, J., Liu, P., & Hou, J. (2006). Fall detection by embedding an accelerometer in cellphone and using KFD algorithm. *Proceedings of international journal of computer science and network security*, 6(10), 277–284.