

Ensemble Learning-Based Algorithms for Aggressive and Agitated Behavior Recognition

Belkacem Chikhaoui^(✉), Bing Ye, and Alex Mihailidis

IATSL Laboratory, Toronto Rehab Institute, University of Toronto, Toronto, Canada
{belkacem.chikhaoui,alex.mihailidis}@utoronto.ca, bing.ye@uhn.ca

Abstract. This paper addresses a practical and challenging problem concerning the recognition of behavioral symptoms dementia (BSD) such as aggressive and agitated behaviors. We propose two new algorithms for the recognition of these behaviors using two different sensors such as a Microsoft Kinect and an Accelerometer sensor. The first algorithm extracts skeleton based features from 3D joint positions data collected by a Kinect sensor, while the second algorithm extracts features from acceleration data collected by a Shimmer accelerometer sensor. Classification is then performed in both algorithms using ensemble learning classifier. We compared the performance of both algorithms in terms of recognition accuracy and processing time. The results obtained, through extensive experiments on a real dataset, showed better performance of the Accelerometer-based algorithm over the Kinect-based algorithm in terms of processing time, and less performance in terms of recognition accuracy. The results also showed how our algorithms outperformed several state of the art methods.

1 Introduction

Globally we are facing a healthcare crisis related to caring for a rapidly aging population who are suffering from a variety of chronic medical conditions, such as dementia. Caring for people with dementia is more complicated given the severity of dementia they suffer from and the degree of autonomy they need for the completion of their activities of daily living [1]. Disruptive behaviors, such as agitation and aggression, are very common in people with dementia and regarded as part of behavioral symptoms of dementia (BSD) [2].

These disruptive behaviors can cause great suffering for persons with dementia and a great deal of distress and burden for caregivers. Early recognition of these behaviors can help effectively provide better treatment for persons with dementia. This, in turn will help reduce caregiver's burden [2]. Direct observation from family caregivers and the care staff is usually used to identify disruptive behaviors. However, this method is subjective, time consuming and could increase the workload of care staff and caregivers [2]. Therefore, researchers

have focused on developing smart systems to automatically monitor and recognize aggression and agitation [3], as these systems will significantly reduce the manpower and time needed to observe and detect these behaviors [4].

Various types of sensors have been used for behavior recognition such as accelerometers, cameras, and Kinects. However, particular attention has been devoted recently to the use of Kinect and the accelerometer sensors given the rich information they provide of a person’s behaviors comparing to other sensors [5]. Kinect, which is a vision sensor, allows collecting different types of data such as individual movements, physical and verbal behaviors in skeleton, depth or color data format. Kinect sensors have been gaining momentum in different domains to monitor people behaviors and considered more accurate than many other 3D sensors [6]. Accelerometer sensor, as opposed to Kinect sensor, is a wearable sensor that can provide acceleration data of person’s behaviors. The acceleration data could be used to develop applications such as recognition of a person’s activity [7] and individual postural tracking [8].

Although Kinect and accelerometer sensors have been used in behavioral recognition, both sensors have not yet been compared for their accuracy in recognition of aggressive and agitated behaviors. In this paper, we use the two sensors to explore their capabilities for the recognition of aggressive and agitated behaviors. We propose two novel algorithms for the recognition of aggressive and agitated behaviors using data collected from a Kinect and an accelerometer sensor. The algorithms extract useful features, which will be combined with an ensemble learning classifier for behavior classification. The primary contributions of this paper can be summarized as follows:

- This work is, to the best of our knowledge, the first formal study that compares Kinect and accelerometer sensors for agitated and aggressive behavior recognition purposes.
- Proposing two novel algorithms for aggressive and agitated behavior recognition using Kinect sensor and accelerometer sensor.
- Collecting real data and conducting extensive experiments to validate the proposed algorithms.

The rest of the paper is organized as follows. First, we give an overview of related work in Sect. 2. Section 3 describes the proposed algorithms in terms of feature extraction, learning and recognition using ensemble method. The results of our experiments on real dataset are presented in Sect. 4. Finally, Sect. 5 presents our conclusions and highlights future work directions.

2 Related Work

Various research studies have been done on normal human action and activity recognition using Kinect and cameras [9,10] and accelerometers [7]. However, very little work has been done on the recognition of aggressive and agitated behaviors using these two sensors. Here we report only research studies interested in aggressive and agitated behavior recognition using Kinect and accelerometer sensors.

2.1 Vision-Based Research Studies

Little work has been done on the recognition of aggressive and agitated behaviors using a Kinect sensor. Nirjon et al. [11] proposed a system to detect aggressive actions such as hitting, kicking, pushing, and throwing from streaming 3D skeleton joint coordinates obtained from a Kinect sensor. The authors combined supervised and unsupervised learning for behavior classification. However, the unsupervised learning used in [11] needs more interventions from the system's users in order to label the behaviors, which is not practical in real settings. Even though their work is similar to our work in terms of aggressive behaviors recognition using skeleton data, the main difference relies on the methodological side in terms of the features used and the classification algorithms employed. In addition, we use two more actions namely the wandering and tearing, which makes our data richer.

Other studies used cameras to monitor [12] and quantify agitation in intensive care unit patients [13]. For example, Chase et al. [13] used patient motion that is sensed using digital video image processing. Then a fuzzy inference system is developed to classify levels of motion that correlate with observed patient agitation. Biswas et al. [14] used multimodal sensors to monitor agitation in people with dementia. The agitation was detected and monitored by the sensors based on the intensity of the movements such as sitting and standing. However, the authors consider only limited movements such as sitting and standing. Siang et al. [4] used a video camera-based method to recognize agitated behaviors. The recorded video data were then annotated based on the gold standard agitation assessment tool to classify agitated behaviors and non-agitated behaviors. Skin color segmentation techniques were used in order to analyse video data and extract relevant features describing agitated behaviors. However, this technique present some limitations in terms of the difficulty in detecting the skin regions during the night and when the person is not facing the camera, which could affect the feature extraction.

2.2 Accelerometer-Based Research Studies

Various studies have been performed on the acceleration of the body joints to understand their association with agitation [15]. However, these studies looked at acceleration data combined with physiological features, such as the galvanic skin response [15], heart rate [16], and skin temperature [16] to train models of agitated behaviors. For instance, Wenhui et al. [15] combined evidences of different modalities to infer stress level. These evidences include physical appearance (facial expression, eye movements, and head movements) extracted from video via visual sensors, physiological conditions (heart rate, skin temperature, Galvanic skin response (GSR) and finger pressure) collected from an emotional mouse, behavioral data from user interaction activities with the computer, and performance measures. Inferring the stress level was modeled using a dynamic Bayesian network. Sakr et al. [16] used wearable sensors to detect agitation. They used bio-physiological measures to detect agitation by monitoring the changes of

the heart rate, galvanic skin response and skin temperature of the participants. Classification is then performed using a support vector machine classifier.

Several other studies [17–19] looked at the correlation between acceleration data and agitated behaviors measured using conventional scales such as Cohen-Mansfield Agitation Inventory (CMAI) and Neuropsychiatric Inventory (NPI) scales. However, no formal methods were proposed in these studies and only correlation measures were investigated. In addition, the weakness of these methods is that, they require an active engagement of the person while collecting data and require physiological data to assess the agitation level, which is obtrusive. In our work, we consider only acceleration data, which can be collected unobtrusively using accelerometers worn on the wrists for example like watches. No physiological data is used in our work.

3 Our Approach

In this section we discuss our proposed algorithms such as the feature extraction and classification using ensemble learning.

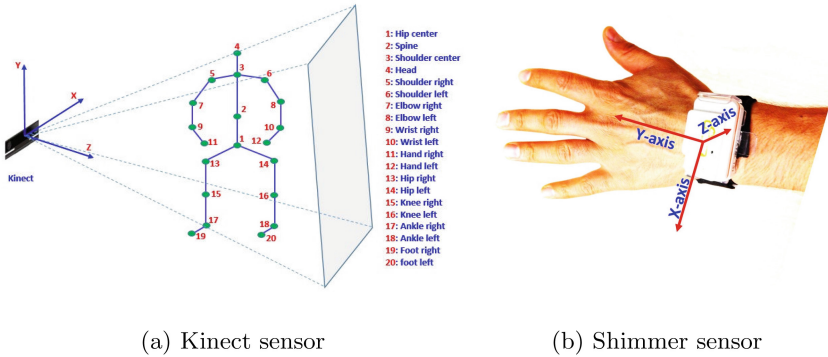


Fig. 1. Kinect and Shimmer sensors used in our experiments

3.1 Kinect-Based Algorithm

This algorithm describes how the recognition of aggressive and agitated behaviors can be achieved using skeleton data collected with a Kinect sensor¹. Figure 1(a) shows the different skeleton joints captured by a Kinect sensor.

A human skeleton can be represented by a hierarchy of joints that are connected with bones. The spatiotemporal features are local descriptions of human motions [20]. Therefore, an action can be described as a collection of time series of 3D positions. The time series of 3D positions represent 3D trajectories of the

¹ <http://www.microsoft.com/en-us/kinectforwindows/purchase/>.

joints in the skeleton hierarchy. In order to accurately represent and differentiate between human actions, taking only 3D positions of the joints and how they evolve over time are not sufficient given the similarity between human actions. In order to obtain a better description and representation of human actions, we extracted the following features:

1. **Relative and absolute joint angles:** they are computed between each two connected limbs, and we represent the skeleton motion data as the changes over time of joint angles computed at each time frame. The aim of computing relative and absolute joint angles is to understand the contribution of each body part in performing actions.
2. **Distance to the hip center:** it represents the distance between the different joints and a fixed point of the skeleton, which is the hip center. The aim of computing this distance is to give more information of the body parts involved in each movement over time.
3. **Distance between current and initial position of joints:** it represents the distance between the position of a joint at time t and its initial position at time t_0 (initial frame). The aim of computing this distance is to characterize the spatial information of each joint, and will further indicate how far the joint will be with respect to its initial position.

These features will be used by the ensemble learning algorithm to perform behavior classification.

3.2 Accelerometer-Based Algorithm

This algorithm describes how the recognition of aggressive and agitated behaviors can be achieved using acceleration data captured using an accelerometer sensor. Figure 1(b) shows the different axis of the accelerometer sensor (Shimmer²) used in our work.

Features were extracted from the filtered accelerometer data using a sliding window (\mathbf{w}). We empirically choose a window size of $\mathbf{w} = 14$ with 50% samples overlapping between consecutive windows. Feature extraction on windows with 50% overlap has demonstrated success in previous work [21]. At a sampling frequency of 50Hz, each window represents data for 0.28 s, which is reasonable given that aggressive actions are usually performed quickly [22]. The window size of 14 yielded better results as well as many training examples. The extracted features are described as follows: Mean, Standard deviation, 25th percentile, 75th percentile, Average Absolute Difference (AAD) [23], Average Resultant Acceleration, Entropy and Covariance. The extracted features take into account the description of the three axis when they are taken separately, two axis conjointly and all the axis together. This will allow to extract rich information about each behavior.

² www.shimmersensing.com.

3.3 Classification

Once the features are computed for skeleton and acceleration data, we can now build a classification model for the aggressive and agitated behaviors. Several classification methods could be used such as SVM, decision trees, and naive Bayes to perform classification. However, these methods have shown to be less accurate when compared to ensemble methods [24]. This motivates us to incorporate ensemble methods to build our classification model. The reason to use ensemble methods is to improve the predictive performance of a given model through combining several learning algorithms. We used rotation forest ensemble method to perform classification.

Rotation forest [25] is an ensemble method proposed to build a classifier based on independently trained decision trees. It is found to be more accurate than bagging, AdaBoost and Random Forest ensembles across a collection of benchmark datasets. The advantage of rotation forests lies in the use of principal component analysis (PCA) to rotate the original feature axes so that different training sets for learning base classifiers can be formed.

Formally, let $\mathbf{x} = [x_1, \dots, x_n]^T$ be a data point described by n features, and let X be an $m \times n$ matrix containing the training example. Let $Y = [y_1, \dots, y_m]^T$ be a vector of class labels for the training data, where y_j takes a value from the class labels $\{w_1, \dots, w_c\}$. Let $D = \{D_1, \dots, D_L\}$ be the ensemble of L classifiers and \mathbf{F} be a feature set. The idea is that all classifiers can be trained in parallel. Therefore, each classifier D_i is trained on a separate training set T_{D_i} to be constructed as follows [25]:

1. split the feature vector \mathbf{F} into K subsets. The subsets may be disjoint or intersecting. Note that rotation forest aims at building accurate and diverse classifiers. Therefore, to maximize the chance of getting high diversity, it is suggested to take disjoint subsets of features. For the sake of simplicity, if K is a factor of n , so that each subset contains $M = n/K$ features.
2. for each of the subsets, select randomly a nonempty subset of classes and then draw a bootstrap sample of objects.
3. run PCA using only the M features in $\mathbf{F}_{i,j}$ and the selected subset of X , where j is the j^{th} subset of features for the training set of classifier D_i . Then, store the obtained coefficients of the principal components $\mathbf{a}_{i,j}^1, \dots, \mathbf{a}_{i,j}^{M_j}$ in a matrix $C_{i,j}$.
4. rearrange the columns of the matrix $C_{i,j}$ in a new matrix B_i^a so that they correspond to the original features in matrix X .
5. the training set for classifier D_i is XB_i^a .
6. to classify a new sample \mathbf{x} , we compute the confidence ψ for each class as follows:

$$\psi_j(\mathbf{x}) = \frac{1}{L} \sum_{i=1}^L d_{i,j}(\mathbf{x}B_i^a), j = 1, \dots, c \quad (1)$$

where $d_{i,j}(\mathbf{x}B_i^a)$ is the probability assigned by the classifier D_i indicating that \mathbf{x} comes from class w_j . Therefore, \mathbf{x} will be assigned to the class having the highest confidence value.

Due to space limitations, the reader is referred to [25] for more details about the rotation forest ensemble method. The next section presents the validation of the two algorithms.

4 Validation

We evaluate the performance of our algorithms on a real human behavior dataset, that contains aggressive and agitated human behaviors obtained by conducting an experiment in Toronto Rehabilitation Institute-UHN (TRI-UHN). Ten (10) participants (6 males and 4 females, 3 among them were left-handed) were involved in this experiment to conduct six (6) aggressive and agitated behaviors (hitting, pushing, throwing, tearing, kicking and wandering) in front of a Kinect sensor v2 by wearing a Shimmer accelerometer sensor. These behaviors have been identified as the most common challenging aggressive and agitated behaviors³ observed from persons with dementia. These behaviors were selected from Cohen-Mansfield Agitation Inventory (CMAI) Scale and are described as follows: **hitting, pushing, throwing, tearing, kicking, and wandering.**

Participants were asked to perform the full set of actions using the right side of the body. For instance, hitting and kicking with the right hand and the right foot respectively. Note that two of these actions, pushing and wandering, are not specific to one side of the body. In order to ensure the study is generic and takes into account both left-handed and right-handed people, participants were then requested to repeat the four laterally specific actions, hitting, kicking, throwing and tearing, using the left side of the body. Participants performed all the actions in front of a Kinect sensor five times while facing each of three directions (front, left and right). For example, during the hitting action, participants did first the action facing the Kinect sensor five times, then repeated the action another five times with their left side facing the Kinect and then another five times with their right side facing the Kinect. This is to ensure that we take into account different situations that might occur when a person is being monitored. A total of 1500 behavior instances have been collected in our experiment from both Kinect and accelerometer sensor. A Research Ethics Board (REB) approval was obtained prior to collecting the data.

Each action was performed using three different directions with respect to the Kinect sensor: front side facing the Kinect, right side facing the Kinect and left side facing the Kinect. Each skeleton data consists of 3D coordinates of 25 joints with time stamp indicating the time when the joint coordinates were recorded at each frame. All the skeleton data were recorded at 30 frame per second rate. The acceleration data were recorded using the Shimmer connect application⁴. All the acceleration data were recorded at 50 Hz frequency. A total of 75 features were extracted from Kinect data and 48 features from the Accelerometer data.

³ Here we use the terms Behavior and Action interchangeably.

⁴ <http://www.shimmersensing.com/shop/shimmer3>.

4.1 Experimental Results

We evaluate the performance of the two proposed algorithms using the TRI dataset. In our experiments, we used the F-Measure to present the results. We used the Leave-One-Out cross validation method to validate our algorithms. In this experiment, we used all behavior instances from 9 participants for training and the behavior instances of the remaining participant for testing. We performed the experiment 10 times, excluding one participant at each time. The benefit of such setup is twofold. First, it allows detecting problematic participants and analyzing the sources of some of the classification errors caused by these participants. A problematic participant means his/her behaviors were performed differently compared to other participants. Second, it allows testing the inter-participant generalization of the algorithms, which constitutes a good indicator about the practicability of our algorithms. For classification performance comparison purposes, we also reported the results using different other single and ensemble classifiers such as random forest, decision tree, Bayesian net, AdaBoost, and Bagging. Table 1 shows the classification results obtained for each algorithm.

Table 1. Recognition results (F-Measure) obtained using the kinect-based and accelerometer-based algorithms

	Kinect-based algorithm		Accelerometer-based algorithm	
	Left handed	Right handed	Left handed	Right handed
Our approach	0.88	0.93	0.85	0.94
Random forest	0.77	0.92	0.81	0.87
Decision tree	0.7	0.83	0.72	0.77
Bayesian net	0.8	0.79	0.45	0.5
AdaBoost	0.44	0.38	0.45	0.51
Bagging	0.72	0.89	0.75	0.83

The results obtained using the Right-handed dataset are promising compared to those obtained using the Left-handed dataset. The good results obtained using the Right-handed dataset can be explained by the fact that the majority of the participants ($n = 7$) were right handed so that behaviors were performed as they normally perform their behaviors. Investigation of the participant errors, in each of the 10 leave one out experiments on the Left-handed dataset, revealed that the most problematic behavior instances belonged to participants number 6 and 7. Indeed, by inspecting the behavior classes with high error rate for participant 7, we found that the participant performed the Hitting behavior by rising the hand behind the head and pretend to hit in exactly the same way as the Throwing behavior, while the other participants punch when performing this behavior without rising their hands behind their head.

Similarly, participant number 6 performed the Throwing behavior with additional movements such as moving left and back while the behavior should be

performed only by hands. Moving left and back when performing the Throwing behavior created confusions with the Wandering behavior where participants were asked to move forward and backward and left and right. Moreover, the participant performed the Tearing behavior by moving the hands forward in the same way as the Pushing behavior, and then performed the Tearing behavior. This creates a confusion with the Pushing behavior. The variability observed in the ways participants performed the different behaviors constitutes a good validation setting for our algorithms. This is demonstrated by the promising results obtained using the Right-handed and the Left-handed datasets.

4.2 Execution Time

The execution time is an essential part in the development of real time applications. We compared the execution time of both algorithms in order to identify which algorithm is more efficient. Although both algorithms use the same ensemble learning classifier, the difference in the execution time will give a good indication about the influence of the number of features used in each algorithm. To measure the execution time, a machine with 6 GB of memory and 2.5 GHz processor is used to perform these experiments. Figure 2 shows the execution time taken by each algorithm to train and classify the data in the Left-handed and Right-handed datasets.

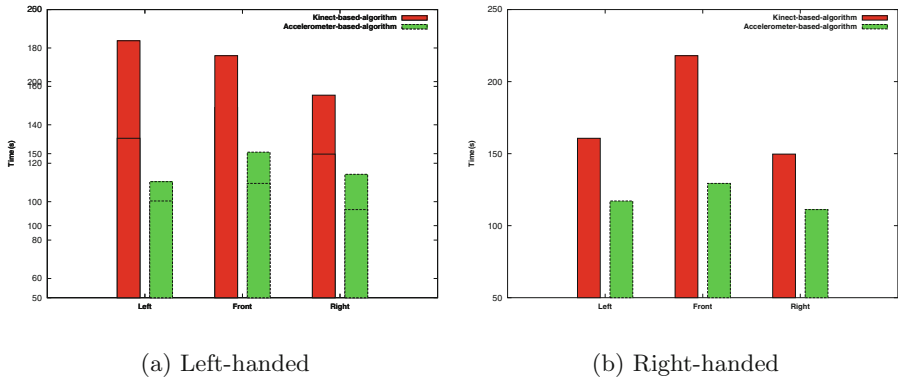


Fig. 2. Execution time taken by each algorithm in the left-handed and right-handed datasets

As shown in Fig. 2, the Accelerometer-based algorithm takes lower time compared to the Kinect-based algorithm in the Left-handed and Right-handed datasets. For instance, the Kinect-based algorithm takes 217.99s to train and classify the Front side data in the Right-handed dataset, while the Accelerometer-based algorithm takes only 129.34s as shown in Fig. 2(a). The same observations were found in the Left-handed dataset. This can be explained by the number of frames the Kinect sensor captured for each behavior. In fact,

data were collected at 30 fps for each behavior, which generates more training samples compared to the acceleration data.

5 Comparison

In order to compare our algorithms with the state of the art methods, we compared the Kinect-based algorithm with the approach of Nirjon et al. [11] for aggressive and agitated behavior recognition using skeleton data collected using a Kinect sensor, and the approach of Zhu et al. [20] for behavior recognition using random forest based classification. Similarly, we compared the Accelerometer-based algorithm with the approach of Bao et al. [21] for human activity recognition using accelerometer data, and the approach of Ravi et al. [7] for activity recognition using acceleration data. The comparison of the Accelerometer-based algorithm with the approaches of [7, 21] is justified by the lack of work reported in the literature about the recognition of aggressive and agitated behaviors using only acceleration data. These two approaches are the most cited works on activity recognition using acceleration data. Table 2 shows the results obtained for each approach.

Table 2. Comparison results for each approach

		Left-handed	Right-handed
Kinect	Kinect-based algorithm	0.88	0.93
	Nirjon et al. [11]	0.31	0.46
	Zhu et al. [20]	0.61	0.65
Accelerometer	Accelerometer-based algorithm	0.85	0.94
	Bao et al. [21]	0.82	0.68
	Ravi et al. [7]	0.78	0.75

As shown in Table 2, both the Kinect-based and Accelerometer-based algorithms outperform the state of the art methods for the Right-handed and Left-handed datasets. Interestingly, the state of the art methods based on acceleration data perform also well compared to our algorithm. Indeed, the method of [21] achieves an accuracy of 0.82 for the Left-handed dataset, while the method of [7] achieves an accuracy of 0.78 for the same dataset. However, the state of the art methods based on the Kinect data achieves low accuracy in both datasets. Therefore, this confirms the effectiveness and suitability of our algorithms.

6 Conclusion

In this paper we have studied the problem of agitated and aggressive behavior recognition. We have proposed and compared two effective algorithms. The

first algorithm used skeleton data collected using a Kinect sensor, while the second algorithm used acceleration data collected using an accelerometer sensor. We have illustrated the effectiveness and suitability of the proposed algorithms through experiments on a real agitated and aggressive behavior dataset. The preliminary results show the suitability of the two algorithms in representing behaviors and distinguishing between them. In addition, we have also illustrated how our algorithms outperformed several of the state-of-the-art methods.

We are currently exploring the fusion of Kinect data and Accelerometer data to improve the recognition accuracy of agitated and aggressive behaviors.

References

1. Mihailidis, A., Boger, J., Craig, T., Hoey, J.: The coach prompting system to assist older adults with dementia through handwashing: an efficacy study. *BMC Geriatr.* **8**(1), 28 (2008)
2. Desai, A., Grossberg, G.: Recognition and management of behavioral disturbances in dementia. *Prim. Care Companion J. Clin. Psychiatry* **3**(3), 93 (2001)
3. Qiu, Q., Foo, S., Wai, A., Pham, V., Maniyeri, J., Biswas, J., Yap, P.: Multimodal information fusion for automated recognition of complex agitation behaviors of dementia patients. In: 2007 10th International Conference on Information Fusion, pp. 1–8. IEEE (2007)
4. Foo, S., Pham, V., Htwe, T., Qiu, Q., Wai, A., Maniyeri, J., Biswas, J., Yap, P.: Automated recognition of complex agitation behavior of demented patient using video camera. In: Conference One-Health Networking, Application and Services, pp. 68–73 (2007)
5. Wouter, V., Egon, V., Reinier, K., Schavemaker, J.: Towards sensing behavior using the kinect. In: 8th International Conference on Methods and Techniques in Behavioural Research, pp. 372–375. Noldus Information Technology (2012)
6. Banos, O., Calatroni, A., Damas, M., Pomares, H., Rojas, I., Sagha, H., Milln, J., Troster, G., Chavarriaga, R., Roggen, D.: Kinect=imu? learning mimo signal mappings to automatically translate activity recognition systems across sensor modalities. In: ISWC, pp. 92–99 (2012)
7. Nishkam, R., Nikhil, D., Preetham, M., Littman, M.: Activity recognition from accelerometer data. *AAAI* **5**, 1541–1546 (2005)
8. Yun, X., Bachmann, E., Moore, H., Calusdian, J.: Self-contained position tracking of human movement using small inertial/magnetic sensor modules. In: IEEE International Conference on Robotics and Automation, pp. 2526–2533. IEEE (2007)
9. Xia, L., Chen, C., Aggarwal, J.: View invariant human action recognition using histograms of 3d joints. In: CVPR Workshops, pp. 20–27 (2012)
10. Chen, C., Jafari, R., Kehtarnavaz, N.: A survey of depth and inertial sensor fusion for human action recognition. *Multimed. Tools Appl.*, 1–21 (2015)
11. Nirjon, S., Greenwood, C., Torres, C., Zhou, S., Stankovic, J., Yoon, H., Ra, H., Basaran, C., Park, T., Son, S. Kintense: a robust, accurate, real-time and evolving system for detecting aggressive actions from streaming 3d skeleton data. In: ACM Conference on Embedded Networked Sensor Systems, pp. 1–9 (2013)
12. Martinez, M., Stiefelhagen, R.: Automated multi-camera system for long term behavioral monitoring in intensive care units. In: Proceedings of the 13 IAPR International Conference on Machine Vision Applications, MVA, pp. 97–100 (2013)

13. Chase, J.G., Agogue, F., Starfinger, C., Lam, Z.H., Shaw, G.M., Rudge, A.D.: Quantifying agitation in sedated ICU patients using digital imaging. *Comput. Method Prog. Biomed.* **76**, 131–141 (2004)
14. Biswas, J., Jayachandran, M., Thang, P., Fook, V., Choo, T., Qiang, Q., Takahashi, S., Jianzhong, E., Feng, C., Yap, P.: Agitation monitoring of persons with dementia based on acoustic sensors, pressure sensors, ultrasound sensors,: a feasibility study. In: *International Conference on Aging, Disability and Independence*, pp. 3–15 (2006)
15. Wenhui, L., Weihong, Z., Zhiwei, Z., Qiang, J.: A real-time human stress monitoring system using dynamic bayesian network. In: *CVPR 2005*, pp. 70–77 (2005)
16. Sakr, G., Elhajj, I., Huijjer, H.: Support vector machines to define and detect agitation transition. *IEEE Trans. Affect. Comput.* **1**(2), 98–108 (2010)
17. Pan, W., Yoshida, S., Liu, Q., Wu, C., Wang, J., Zhu, J., Cai, D.: Quantitative evaluation of severity of behavioral and psychological symptoms of dementia in patients with vascular dementia. *Transl. Neurodegener.* **2**(9), 2–7 (2013)
18. Tractenberg, R., Singer, C., Cummings, J., Thal, L.: The sleep disorders inventory: an instrument for studies of sleep disturbance in persons with alzheimer’s disease. *J. Sleep Res.* **12**(4), 331–337 (2003)
19. Knuff, A.: Application of actigraphy to the measurement of neuropsychiatric symptoms of agitation in dementia. Master’s thesis, Queen’s University, Canada (2014)
20. Zhu, Y., Wenbin, C., Guodong, G.: Fusing spatiotemporal features and joints for 3d action recognition. In: *CVPRW*, pp. 486–491 (2013)
21. 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)
22. Stern, T.: *Massachusetts General Hospital Handbook of General Hospital Psychiatry*, 6th edn. Saunders, Elsevier, New York (2010)
23. Kwapisz, J., Weiss, G., Moore, S.: Activity recognition using cell phone accelerometers. *SIGKDD Explor. Newsl.* **12**(2), 74–82 (2011)
24. Opitz, D., Maclin, R.: Popular ensemble methods: an empirical study. *J. Artif. Intell. Res.* **11**, 169–198 (1999)
25. Rodriguez, J., Kuncheva, L., Alonso, C.: Rotation forest: a new classifier ensemble method. *IEEE Trans. Pattern Analy. Mach. Intell.* **28**(10), 1619–1630 (2006)