A New Kinect Approach to Judge Unhealthy Sitting Posture Based on Neck Angle and Torso Angle

Leiyue Yao, Weidong Min[✉], and Hao Cui

School of Information Engineering, Nanchang University, Nanchang 330031, China
minweidong@ncu.edu.cn

Abstract. Sitting posture has a close relationship with our health, keeping right sitting posture is important for people to avoid chronic diseases. However, automatic unhealthy sitting posture detection system is rare, especially for those based on computer vision technology. This paper proposes a new method of judging unhealthy sitting posture based on neck angle and torso angle detection using Kinect sensor. The method tracks neck angle and torso angle as two representative features from the depth image in a given period of time to judge whether the sitting posture is healthy or not. Experimental results show that the proposed method can judge sitting posture effectively for different unhealthy sitting types. Compared with the existing methods of action recognition, our method only needs a Kinect sensor without any other wearable sensors and is time efficient and robust because of only calculating two angles.

Keywords: Sitting posture judgment · Neck angle · Torso angle · Depth image

1 Introduction

Sitting is one of the most common postures in daily activities. Unhealthy sitting postures will inevitably increase the risk of musculoskeletal disorders [1] while good sitting posture is helpful for children growth and eyesight protection. For these and many more other reasons, the number of researchers on sitting posture correction is getting larger and larger over recent years. With the development of computer science and electronic sensors, automatic unhealthy sitting posture detection methods give help for us to form a good sitting habit. In general, these methods can be divided into two types, methods based on wearable devices and methods based on computer vision.

Mattmann et al. [2] researched on recognizing 27 upper body postures using a garment with strain sensors. The method of Harms et al. [3] recognized 21 human exercise postures through a smart shirt system (SMASH) with acceleration sensors. Karantonis et al. [4] used a waist-mounted tri-axial accelerometer system to classify human movement status. The approach of Jeong et al. [5] used a 3-axis accelerometer to monitor human's activity volume and recognize emergent situations. Barba et al. [6] aimed at creating a sensor capable of providing detection measures at the least possible cost. Although the above wearable sensors approaches gathered sufficient information to detect, classify and recognize human activities and worked well in posture detection and recognition, they have some obvious disadvantages. The wearable systems can be a

source of inconvenience or discomfort. Moreover, the wearable devices can be worn gradually or damaged due to external factors such as being squeezed or pressed, which lead to failure to collect information.

Due to these main disadvantages of the wearable device based approaches, the methods based on computer vision become a hot topic in recent researches. The computer vision based methods need no wearable devices and extract features from videos using image processing technologies. Li and Chen [7] researched on recognizing human postures including standing, sitting, kneeling and stooping through analyzing 10 parameters extracted from video frames. Boulay et al. [8, 9] proposed a 3D human-body-posture recognition method to recognize standing, sitting, lying down and stooping by comparing horizontal and vertical projections of human body with their corresponding predefined 3D human posture models. A method proposed by Wang et al. [10] used background subtraction on the depth image created by Kinect sensor to extract a silhouette contour of a human, and determined different type of activities using a pre-trained LVQ (Learning Vector Quantization) neural network.

Among the computer vision methods, approaches based on machine learning, deep learning and neural network are popular and work well in posture recognition. Althloothi et al. [11] presented a shape representation and kinematic structure, and used the MKL (Multiple Kernel Learning) technique at the kernel level for human activity recognition. Ruizhi and Lingqiao [12] achieved posture estimates by encoding each local descriptor, named as trajectorylet in their method, using a discriminative trajectorylet detector set which is selected from a large number of candidate detectors trained through exemplar-SVMs. Because of environmental problems and intrinsic noise, videos of similar actions may suffer from huge intra-class variations. Jalal et al. [13] solved these problems by introducing the ELS-TSVM (Energy-based Least Square Twin Support Vector Machine) algorithm. Jiayu et al. [14] proposed an abstract and efficient motion tensor decomposition approach to compress and reorganize the motion data. Together with a multi-classification algorithm, the approach is able to efficiently and accurately differentiate various postures. Vina and Mohamad [15] proposed a distribution-sensitive learning method based on RVM (Relevance Vector Machine) to recognize pose-based human gesture and solve imbalanced data problem. The gesture recognition method of Liu et al. [16] constructed a 3D2CNN (3D-based Deep Convolutional Neural Network) to directly learn spatio-temporal features, and then computed a joint based feature vector named JointVector for each sequence using the simple position and angle information between skeleton joints. Li et al. [17] proposed a feature learning approach based on SAE (Sparse Auto-Encoder) and principle component analysis for recognizing human gestures. Leng et al. [18] proposed a novel 3D model recognition mechanism based on DBM, which can be divided into two parts, the feature detection based on DBM and the classification based on semi-supervised learning method.

Since the depth image provides more information than color image, more and more researchers pay special attention to detect and recognize human postures and activities through analyzing the depth image. Kinect sensor is widely used because of its competitive price and exceptional performance in skeleton tracking. Ibañez et al. [19] proposed a lightweight approach to recognize gestures with Kinect through utilizing approximate string matching. In literatures [20–23], several other methods based on the depth image

and Kinect sensors have been proposed for fall detection. Manghisi et al. [24] developed a semi-automatic evaluation software based on Kinect V2 to detect awkward postures in real time. Since depth image is unaffected by illumination of the environment and shadows, Wang et al. [25] proposed a new method based on depth image for human shape object detection and a combined method to recognize five postures. In literatures [26], spatiotemporal features created by RGB-D video sequences are used for human tracking and activity recognition. Alwani et al. [27] calculated joint angles by using the skeleton information to describe the sequences of actions. A Hidden Markov Model was used to classify actions. With the help of skeleton data and 3D joint positions, there are lots of new method emerged in the field of action recognition and classification. However, as for judging whether an action is correct or not, few attentions are paid to this research field.

As discussed above in details, the wearable based approaches have some limitations, while the deep learning based approaches based on computer vision are time consuming and depend on training data sets. To alleviate these issues, this paper presents a new method based on neck angle and torso angle detection in depth images to effectively recognize unhealthy sitting posture. The neck angle and the torso angle are the most representative two features of unhealthy sitting posture [28], therefore they are adopted as the criteria of sitting posture judgment in our approach. Being different from the wearable device based approaches and the deep learning based approaches, our approach only need a Kinect sensor without any other wearable sensors and is time efficient and robust because of only calculating two angles. Moreover, only usage of depth image ensures privacy protecting in some senses.

The following parts of this paper are organized as follows. Section 2 describes our proposed method of unhealthy sitting posture judgment. Section 3 discussed the experiments and results of our approach. Conclusions are given in Sect. 4.

2   Our New Approach for Unhealthy Sitting Posture Detection

2.1   Sitting Posture Modeling and Feature Extraction

Before judging the unhealthy sitting posture, criterions which can be used to differentiated healthy sitting posture and unhealthy sitting posture should be built first. In this paper, PEO (Portable Ergonomic Observation Method) [29] is adopted to model sitting posture. Thanks to the clear definition of unhealthy sitting posture in PEO, two representative features, i.e. the neck angle and the torso angle, can be extracted from the model. Figure 1 shows the comparison of healthy and unhealthy sitting gesture.

Figure 1(a) shows the typical healthy sitting gesture where two angles, the neck angle and the torso angle, are taken into consideration regarding whether the upper body and the head line up with the gravity direction, respectively. The neck angle is the angle between the vector from the head to the neck and the gravity vector. The torso angle is the angle between the vector from the neck to the spine base and the gravity vector. Unhealthy sitting gesture can be identified when the value of either of these two angles is larger than a given threshold 20°. McAtamney et al. [28] proved that the sitting posture
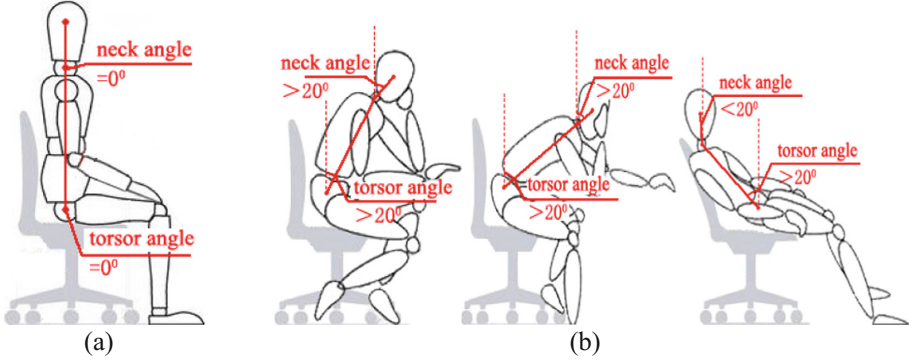
Fig. 1. Sitting posture model. (a) Healthy sitting gesture. (b) Three typical unhealthy sitting gestures.

is unhealthy when the torso angle is over 20°. Figure 1(b) shows three typical unhealthy sitting gestures.

Since sitting, especially sitting for work, is an activity occupying a long period of time, unhealthy sitting posture cannot be simply judged by one or two frames containing abnormal postures. In our proposed method, we take a certain period of time as a unit of measurement and calculate the ratio of unhealthy frames to total frames. When the ratio is larger than a given threshold, it is detected as the unhealthy sitting posture. Detailed algorithm is elaborated in Sect. 2.3.

2.2 Key Joints Acquisition and Angles Calculation

It is a hard work to estimate the joints of one's body in a RGB image. However, the depth image created by Kinect 2.0 sensor provides not only 3D information but also 25 joints information of a person when the person stands and 10 joints information when the person sits.

Every joint contains three-dimensional information and its tracking state. Although the spin joint cannot be tracked, it can be estimated by Kinect program. To calculate the neck angle and the torso angle discussed in Sect. 2.1, three joints of head, shoulder-center and spin, are adopted to fulfill this task. Formulas (1) and (2) are used to calculate the following two vectors. One is the vector from head to shoulder-center ($\overrightarrow{HSc}$). Another vector is from shoulder-center to spin ($\overrightarrow{ScSp}$).

$$\overrightarrow{HSc} = (X_h - X_{sc}, Y_h - Y_{sc}, Z_h - Z_{sc}) \tag{1}$$

$$\overrightarrow{ScSp} = (X_{sc} - X_{sp}, Y_{sc} - Y_{sp}, Z_{sc} - Z_{sp}) \tag{2}$$

Here, $H(X_h, Y_h, Z_h)$, $Sc(Xs_c, Ys_c, Zs_c)$ and $Sp(Xs_p, Ys_p, Zs_p)$ mean the position of head, shoulder-center and spin in three-dimensional space, respectively.

As shown in Fig. 2(a), both angles are formed through the gravity vector. We take the neck angle as an example to explain the way we calculate its value. Figure 2 shows the detailed steps.
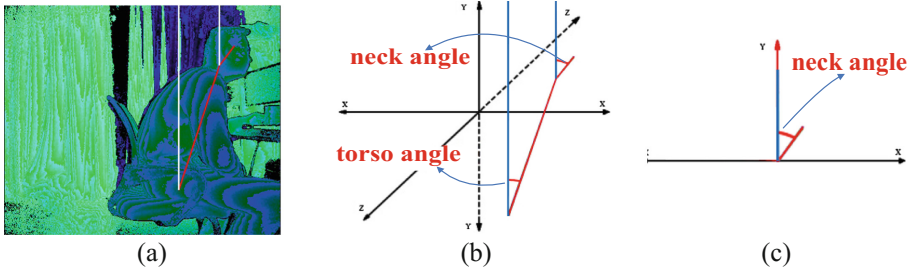

Fig. 2. Extracting the neck angle and the torso angle features from a depth image. (a) Depth image of a typical unhealthy sitting posture. (b) Vectors and angles in 3D space. (c) neck angle transformed in 2D plane

Firstly, once a person is detected, vector $\overrightarrow{HSc}$ can be extracted from the frame. Figure 2(a) shows the vector line which is calculated and drawn in depth image. Figure 2(b) is its presentation in 3D coordinate.

Secondly, according to vector translation theory in solid geometry, point Sc in $\overrightarrow{HSc}$ can be moved to the origin coordinate, and the vector can be presented in 2D coordinate. Figure 2(c) shows its final status.

Thirdly, since the gravity vector is always vertical to the ground, any point on y-axis and the origin coordinate can be used to form the gravity vector. To keep gravity vector and $\overrightarrow{HSc}$ in the same 2D plane, the position of gravity point in 3D coordinate can be defined as $G(Xsc, 0, Zxc)$, where the values in x-axis and z-axis are the same as shoulder-center. The gravity vector can be calculated using formula (3).

$$\overrightarrow{GSc} = (Xg - X_{Sc}, Yg - Y_{Sc}, Zg - Z_{Sc}) \tag{3}$$

Since Xg and Zg are the same as Xsc and Zxc, formula can also be simplified as follows in formula (4).

$$\overrightarrow{GSc} = (0, -Y_{Sc}, 0) \tag{4}$$

Forthly, when two vectors have been defined, the neck angle can be calculated using formula (5).

$$\cos(\alpha) = \cos(\overrightarrow{HSc}, \overrightarrow{GSc}) = \frac{\overrightarrow{GSc} \bullet \overrightarrow{HSc}}{|\overrightarrow{GSc}| \times |\overrightarrow{HSc}|} \tag{5}$$

These 4 steps elaborated above can also be used to calculate the torso angle in the same way.

2.3 Algorithm and Implementation

In our experiments, we used Kinect with 30 frames per second. Sitting is such an activity that does not change dramatically in a certain period of time. Therefore, there is no need to analyze every frame and calculate the angles. To reduce computational load, we extract and analyze one frame per 100 ms in our proposed method. Figure 3 shows the general block diagram.
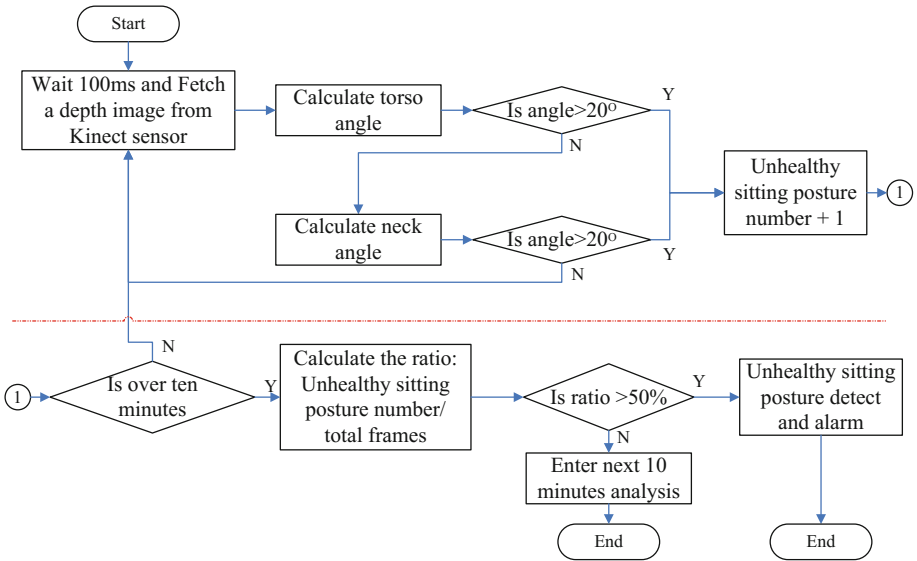

Fig. 3. General block diagram of our approach (Color figure online)

The general block can be divided into two parts. The upper portion above the red line in Fig. 3 shows the key steps of extracting the two angles in our algorithm. Once a person comes into the scene, the program starts to track the person. Every other 100 ms, a depth image will be extracted for neck angle and torso angle calculating. The number of total frame adds 1 and unhealthy sitting posture frames add 1 only when neck angle or torso angle is larger than 20°. Compare with neck, torso shows more times in unhealthy sitting posture, therefore torso angle is first calculated, while neck angle is calculated only when torso angle is less than 20°. In most cases, this order can reduce half computational load.

The lower portion under the red line in Fig. 3 shows the key steps of applying the criteria of unhealthy sitting posture judgment. Since sitting posture is an activity occupying a long period of time, the unhealthy posture is judged every 10 min. That is to say, every 10 min the UFR (Unhealthy Frame Ratio, i.e., the ratio of unhealthy frames to the total frames) is calculated and compare with the given threshold. If the ratio is larger than a given threshold 50%, the 10 min sitting posture is judged as unhealthy, or else it is healthy.

3   Experiment and Results

Our method was implemented using VisualStudio2013 + emgu.cv3.1 + Kinect on a PC using an Intel Core i7-4790 3.60 GHz processor, 8 GB RAM clocked at 1.333 GHZ. All tests were captured from a Kinect 2.0 sensor in BMP format of 640×480 resolution. The tests consist of 5 different types of sitting postures which include 3 healthy sitting posture videos and 30 unhealthy sitting posture videos in 4 different types. Since 10 min sitting posture supervision is too long, which can be compressed and simulated in a shorter period of time without losing generality and correctness of our experiments. All videos in our experiments are kept in 1 min for saving experimental time.

Before calculating neck angle and torso angle, the pre-work which should be guaranteed is finding out the position of three joints in depth image. Figure 4 shows that our proposed method is able to find out the exact position of head, shoulder-center and spin. However, the Kinect sensor should be installed properly because installing in a too high or too low position will have strong effect on the accuracy of joints recognition.
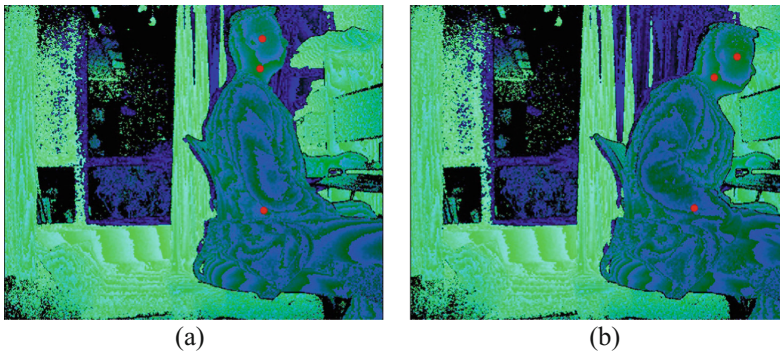

(a)   (b)

Fig. 4.   Three joints in one's skeleton when sitting. (a) healthy sitting posture; (b) unhealthy sitting posture.

Since sitting is a long time activity. Single frame or only a few frames containing unhealthy sitting posture cannot be the basis of judgment. For example, stretching and twisting are common activities during sitting. Therefore, we take the proportion of unhealthy sitting posture frames as the criterion. Figure 5 shows the angles' change curves of a typical unhealthy sitting posture in 1 min video, which contains 201 frames in total, in where there are 130 unhealthy frames and 71 healthy frames. The UFR reaches 61.9%.
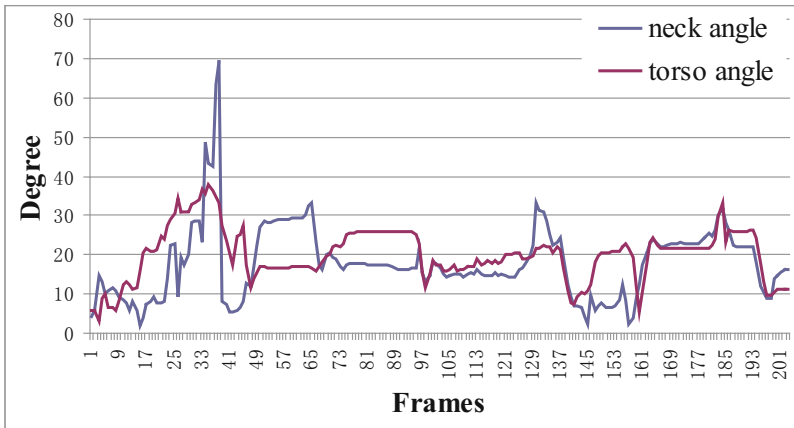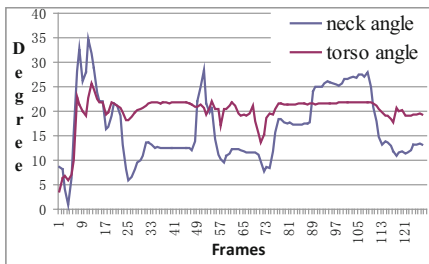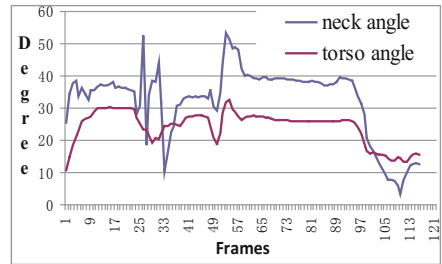
Fig. 5. The angles' change curve of a typical unhealthy sitting posture video.
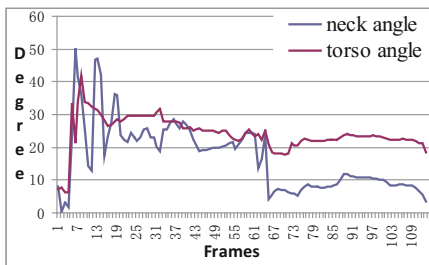
For most situation, unhealthy sitting posture always accompany with excessive tilt torso. Figure 6 shows angles' change curves in 4 typical unhealthy sitting postures.
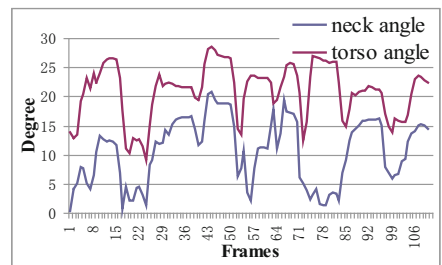

(a)

(b)

(c)

(d)

Fig. 6. Typical unhealthy sitting postures. (a) Reading with head down (UFR: 68.6%). (b) Right leaning on the chair (UFR: 84.6%). (c) Sprawling in the chair (UFR: 89.5%). (d) Sitting with body moving right to left (UFR: 68.5%).

According to the line charts in Fig. 6, it can be inferred that for most situations neck angle is not necessary to be calculated. Because when neck angle bends over 20°, torso angle always reaches its given threshold. Therefore, it is an efficient way to reduce the computational load by calculating torso angle first.

Comparing our method with wearable device approach and deep learning approach, our approach has advantages in real-time calculation and robustness because of its low cost for computing representative features. Our experiment results indicate that the method effectively detected unhealthy sitting postures, and distinguished healthy sitting postures from unhealthy sitting postures. Table 1 shows the detailed information of the experiments.

Table 1. Statistics for sitting posture judgments

Sitting posture types	Sample numbers	Correct identification of sitting posture	Accuracy
Healthy sitting posture	3	3	100%
Lean against the chair	5	5	100%
Lean forward	5	5	100%
Lean backward	5	5	100%
Sitting with body moving	15	11	73.3%

Only 4 unhealthy sitting postures of the type "sitting with body moving" could not be detected correctly. Upon reviewing the records of these 4 videos, we found out that the person in these 4 test videos changed his postures too frequently to accord with human common sitting posture. Therefore, although the accuracy of "sitting with body moving" is only 73.3%, it can be anticipated that our method will work even better in the practical use.

4   Conclusions

In this paper, we proposed a new and fast method of unhealthy sitting posture judgment which is based on the neck angle and the torso angle using depth image captured from a Kinect sensor. In our method only two angles need to be calculated, so it is robust and of high time efficiency. The torso angle is calculated first and used in the judgment. In most situations, when the neck angle reaches the given threshold, it is usually accompanied with a high value of torso angle. Judging sitting posture in this way can further reduce computational load. Experimental results show that the proposed method can judge sitting posture effectively for different unhealthy sitting types. Compared with the existing wearable device based approaches and the deep learning based approaches, our method only needs a Kinect sensor without any other wearable sensors and is time efficient and robust because of only calculating two angles. What is more, our method is based on published medical findings of unhealthy sitting posture judgment condition, and therefore having solid theoretical foundation.

In the future work, we will investigate using deep learning methods to further improve the accuracy of our current method. Moreover, the unhealthy sitting posture

judgment criteria and our proposed algorithm may be extended to other methods using other types of video cameras such as the more common-used monocular camera.

References

1. Lis, A.M., Black, K.M., Korn, H., Nordin, M.: Association between sitting and occupational LBP. Eur. Spine J. **16**(2), 283–298 (2007)
2. Mattmann, C., Amft, O., Harms, H., Clemens, F., Clemens, F.: Recognizing upper body postures using textile strain sensors. In: 11th IEEE International Symposium on Wearable Computers, MA, Boston, pp. 29–36 (2007)
3. Harms, H., Amft, O., Roggen, D., Troster, G.: Rapid prototyping of smart garments for activity-aware applications. J. Ambient Intell. Smart Environ. **1**(2), 87–101 (2009)
4. Karantonis, D.M., Narayanan, M.R., Mathie, M., Lovell, N.H., Celler, B.G.: Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring. IEEE Trans. Inf Technol. Biomed. **10**(1), 156–167 (2006)
5. Jeong, D.U., Kim, S.J., Chung, W.Y.: Classification of posture and movement using a 3-axis accelerometer. In: International Conference on Convergence Information Technology, pp. 837–844. IEEE (2007)
6. Barba, R., de Madrid, Á.P., Boticario, J.G.: Development of an inexpensive sensor network for recognition of sitting posture. Int. J. Distrib. Sens. Netw. **11**(8), 1–13 (2015)
7. Li, C.C., Chen, Y.Y.: Human posture recognition by simple rules. In: IEEE International Conference on Systems, Man and Cybernetics, pp. 3237–3240 (2007)
8. Boulay, B., Bremond, F., Thonnat, M.: Posture recognition with a 3D human model. In: The IEE International Symposium on Imaging for Crime Detection and Prevention, pp. 135–138. IET (2005)
9. Boulay, B., Brémond, F., Thonnat, M.: Applying 3D human model in a posture recognition system. Pattern Recogn. Lett. **27**(7), 1788–1796 (2006)
10. Wang, W.J., Chang, J.W., Haung, S.F., Wang, R.J.: Human posture recognition based on images captured by the kinect sensor. Int. J. Adv. Robot. Syst. **13**(2), 1–16 (2016)
11. Althloothi, S., Mahoor, M.H., Zhang, X., Voyles, R.M.: Human activity recognition using multi-features and multiple kernel learning. Pattern Recogn. **47**(5), 1800–1812 (2014)
12. Qiao, R., Liu, L., Shen, C., Hengel, A.V.D.: Learning discriminative trajectorylet detector sets for accurate skeleton- MARK based action recognition. Pattern Recogn. **66**(6), 202–212 (2017)
13. Nasiri, J.A., Charkari, N.M., Mozafari, K.: Energy-based model of least squares twin support vector machines for human action recognition. Sig. Process. **104**(6), 248–257 (2014)
14. Chen, J., Qiu, J., Ahn, C.: Construction worker's awkward posture recognition through supervised motion tensor decomposition. Autom. Constr. **77**(1), 67–81 (2017)
15. Ayumi, V., Fanany, M.I.: Distribution-sensitive learning on relevance vector machine for pose-based human gesture recognition. In: The Third Information Systems International Conference, pp. 527–534 (2015)
16. Liu, Z., Zhang, C., Tian, Y.: 3D-based deep convolutional neural network for action recognition with depth sequences. Image Vis. Comput. **55**(11), 93–100 (2016)

17. Li, S.Z., Yu, B., Wu, W., Su, S.Z., Ji, R.R.: Feature learning based on SAE–PCA network for human gesture recognition in RGBD images. Neurocomputing **151**(5), 565–573 (2015)
18. Leng, B., Zhang, X., Yao, M., Xiong, Z.: A 3D model recognition mechanism based on deep Boltzmann machines. Neurocomputing **151**(3), 593–602 (2015)
19. Ibañez, R., Álvaro, S., Teyseyre, A., Rodriguez, G., Campo, M.: Approximate string matching: a lightweight approach to recognize gestures with Kinect. Pattern Recogn. **62**(2), 73–86 (2017)
20. Mastorakis, G., Makris, D.: Fall detection system using Kinect's infrared sensor. Real-Time Image Proc. **9**(4), 635–646 (2014)
21. Rougier, C., Auvinet, E., Rousseau, J., Mignotte, M., Meunier, J.: Fall detection from depth map video sequences. In: Abdulrazak, B., Giroux, S., Bouchard, B., Pigot, H., Mokhtari, M. (eds.) ICOST 2011. LNCS, vol. 6719, pp. 121–128. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-21535-3_16
22. Kwolek, B., Kepski, M.: Fuzzy inference-based fall detection using Kinect and body-worn accelerometer. Appl. Soft Comput. **40**(3), 305–318 (2016)
23. Gasparrini, S., Cippitelli, E., Spinsante, S., Gambi, E.: A depth-based fall detection system using a Kinect® sensor. Sensors **14**(2), 2756–2775 (2014)
24. Manghisi, V.M., Uva, A.E., Fiorentino, M., Bevilacqua, V., Trotta, G.F., Monno, G.: Real time RULA assessment using Kinect v2 sensor. Appl. Ergon. **2**(1), 1–11 (2017)
25. Wang, W.J., Chang, J.W., Haung, S.F., Wang, R.J.: Human posture recognition based on images captured by the Kinect sensor. Int. J. Adv. Rob. Syst. **13**(1), 1–16 (2016)
26. Kamal, S., Jalal, A.: A hybrid feature extraction approach for human detection, tracking and activity recognition using depth sensors. Arab. J. Sci. Eng. **41**(3), 1043–1051 (2016)
27. Alwani, A.A., Chahir, Y., Goumidi, D.E., Molina, M., Jouen, F.: 3D-posture recognition using joint angle representation. In: Laurent, A., Strauss, O., Bouchon-Meunier, B., Yager, R.R. (eds.) IPMU 2014. CCIS, vol. 443, pp. 106–115. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-08855-6_12
28. Mcatamney, L., Corlett, E.N.: RULA: a survey method for the investigation of work-related upper limb disorders. Appl. Ergon. **24**(2), 91–99 (1993)
29. Fransson-Hall, C., Gloria, R., Kilbom, A., Winkel, J., Karlqvist, L., Wiktorin, C.: A portable ergonomic observation method (PEO) for computerized on-line recording of postures and manual handling. Appl. Ergon. **26**(2), 93–100 (1995)