Development of Human Fall Detection System using Joint Height, Joint Velocity and Joint Position from Depth Maps

Yoosuf Nizam[2], Mohd Norzali Haji Mohd[1], Razali Tomari[3], M. Mahadi Abdul Jamil[2]
[1]Embedded Computing Systems (EmbCos),
Department of Computer Engineering
[2]Modeling and Simulation (BIOMEMS) Research Group,
Department of Electronic Engineering
[3]Advanced Mechatronic Research Group (ADMIRE),
Department of Mechatronic and Robotic Engineering,
Faculty of Electrical and Electronic Engineering,
Universiti Tun Hussein Onn Malaysia,
86400 Parit Raja, Batu Pahat, Johor, Malaysia.
hpa7890@gmail.com

***Abstract*—Human falls are a major health concern in many communities in today's aging population. There are different approaches used in developing fall detection system such as some sort of wearable, ambient sensor and vision based systems. This paper proposes a vision based human fall detection system using Kinect for Windows. The generated depth stream from the sensor is used in the proposed algorithm to differentiate human fall from other activities based on human Joint height, joint velocity and joint positions. From the experimental results our system was able to achieve an average accuracy of 96.55% with a sensitivity of 100% and specificity of 95%.**

***Index Terms*—Kinect; Velocity; Joint Height; Depth Maps.**

I. INTRODUCTION

Fall detection for elderly is a major topic as far as assistive technologies are concerned. Human fall detection systems are important, because fall is the main obstacle for elderly people to live independently. Moreover statistics [1, 2] has also shown that falls are the main reasons of injury related deaths for seniors aged 79. Fall detection systems use various approaches to distinguish human fall from other activities of daily life, such as wearable based devices, non-wearable sensors and vision based devices.

The two most common approaches used in human fall detection systems are wearable and ambient based methods. However, due to the high false alarm ratio most of the users and even medical care givers are not ready to accept and depend on such devices. Therefore recent researches are focused on non-wearable (vision) based approaches for human fall detection. Vision based approaches using live cameras are accurate in detecting human falls and it does not generate high false alarms. However normal cameras require adequate lighting and it does not preserve the privacy of users. As a result users are not willing to accept having such systems installed at their home. Moreover, with normal RGB camera it is not possible to achieve the accuracy than that of a depth sensor, in extracting the position of the subject. One of the sensors that generate depth images to track human skeleton is Kinect for windows. This paper proposes a vision based fall detection system, which uses Kinect sensor to capture depth image.

II. PREVIOUS WORKS

The manuscript article should be written in English in the font of Times New Roman, which includes the following: abstract, introduction, literature review, objectives, research methodology, theory, testing and analysis, results and discussions, conclusion, acknowledgement and references. Manuscript should be prepared via the Microsoft Word processor. There are basically three approaches used to develop human fall detection systems as wearable based device, camera (vision) based and ambience based devices. Among them Vision based fall detection systems are more popular due to the fact that it can identify and classify human movements accurately. Thus, vision based sensors [3, 4] for human detection and identification are important sensors among the researchers, especially as they tend to base their fall detection on non-wearable sensors [5]. The depth cameras, classified in vision based sensors are more accurate than normal RGB cameras in human detection. Since this paper focuses on the use of the depth information to recognize fall, we will review only a set of selected papers that had based their fall detection on depth sensors.

Kepski and kwolek used a Kinect sensor as a ceiling mounted depth camera for capturing the depth images [6]. Than they used k-NN classifier to separate lying pose from normal daily activities and applies distance between head to floor to identify fall. In order to distinguish between intentional lying postures and accidental fall they also used motion between static postures.

Combination of a wearable wireless accelerometer and a depth sensor based fall detection were conducted in [7, 8] which used distance between the person centre of gravity and floor to confirm fall. The authentication of fall after potential fall indication from the accelerometer is accomplished from a Kinect sensor depth images.

Yang et al proposed a robust method based on Spatio temporal context (STC) tracking of depth images from a Kinect sensor. In the pre-processing, the parameters of the single Gauss Model (SGM) are estimated and the coefficients of the floor plane are extracted. Foreground coefficient of ellipses is used to determine the head position and STC algorithm is used to track head position. The distance from head to floor plane is calculated in every following frame and a fall indicated if an adaptive threshold is reached [9].

Bian et al presented a method for fall detection based on two features: distance between human skeleton joints and the floor, and the joint velocity. A fall is detected if the distance between the joints and the floor is close. Then the velocity of the joint hitting the floor is used to distinguish the event from a fall accident or a sitting/lying down on the floor [5].

A fall detection and reporting system using Microsoft Kinect sensor presented in [10], uses two algorithms. The first uses only a single frame to determine a fall and the second uses time series data to distinguish between fall and slow lying down on the floor. For these algorithms they use the joint position and the joint velocities. The reporting can be sent as emails or text messages and can include pictures during & after the fall.

Gasparrini et al proposed an automatic, privacy-preserving, fall detection method for indoor that uses Microsoft Kinect sensor on ceiling configuration. Ad-Hoc segmentation algorithm is used recognize the elements captured in the depth scene. Then blobs in the scene are classified and anthropometric relationships and features are exploited to recognize one or more human subjects among the blobs. Once a person is detected, he is followed by a tracking algorithm and a fall is detected if the blob associated to a person is near to the floor [11].

In another study, a mobile robot system is introduced, which follows a person and detects when the person has fallen using a Kinect sensor. They used the distance between the body joints and the floor plane to detect fall [12].

Rougier et al presented a method for fall detection that uses human centroid height relative to the ground and body velocity. They have also dealt with occlusions, which was a weakness of previous works and claimed to have a really good fall detection results with an overall success rate of 98.7% [13].

The related works either uses the distance from head to floor and the velocity of the joints or position and velocity of joints to classify human falls. They used Different algorithms to extract the subject from the depth images. The algorithms used to classify human falls were tested on simulated falls and other activities. The proposed algorithm uses combination human joint velocity and height to differentiate human fall from other activities and then the position of the joints are used to confirm a fall.

III. METHODOLOGY

In our methodology we use the depth information generated from Kinect v1 sensor for 3D human joint position and floor plane extraction. The depth images are generated from the IR sensor stream, which can work both day and night. The Kinect sensor used here, can detect 3D location of 21 joints for two people in 'default' mode (10 joints in 'seated' mode) using Kinect SDK. This capability to track the skeleton image of one or two people within the Kinect's field of view is processed by the Kinect Runtime from the generated depth images. These data are then used to compute the velocity of the body, the distance between the head to floor plane and the position of the other joints to identify an unintentional fall.

Unlike the other related works, fall detection in the proposed system is accomplished using both the velocity of head, the distance from head to floor and position of other joints. The system will be continuously sensing for changes in the velocity of head joints and distance from head to floor. If the system notices for any abnormal change it will check the position of the joints to confirm a fall activity. The system will confirm a fall if the subject remains on the floor without any movement for 5 seconds and then a fall alarm will be generated. The developed algorithm was tested on simulated falls and other Activities of daily life to generate necessary results for comparing the performance with the related works. The number of test trials for fall from standing and fall from chair are 17 and 19 trials respectively. The total test includes 9 brutal movements and 35 trials including activities that may lead to or are similar to potential fall activity. As a comparison, the use of the joint positions to confirm human fall after fall detection from joint height and joint velocity helped to reduce the false alarm. At the same time these of all joints for identifying the position of subject by ignoring or avoiding the joints that are not properly detected can make sure that the system will confirm a fall nearly in all conditions.

A. *Floor Plane and Human detection*

The detected skeleton joints data are stored as (x, y, z) coordinates, which is expressed in meters. In this coordinate system, the positive y-axis extends vertical upwards from the depth sensor, the positive x-axis extends to the left placing the Kinect sensor on a surface level and the positive z-axis extending in the direction in which the Kinect is pointed. The z value gives the distance of the object to the sensor (objects close to sensor will have a small depth value and object far away will have larger depth value). Using these joint coordinate data, the movement of any joint, velocity can be computed along with the direction respective to the previous joint position.

To calculate the distance between the joints to floor, we need to find the floor plane. In our system we use the floor plane provided by the Kinect. To find out the floor plane equation at start up, three points from 3D floor plane is chosen and then the system automatically solve the floor plane equation. The skeleton frame generated from depth image also contains floor-clipping-plane vector, which has the coefficients of an estimated floor-plane equation as shown by Equation (1). This clipping plane was used for removing the background and segmenting players.

$$Ax + By + Cz + D = 0 \quad (1)$$

where: A = vFloorClipPlane.x
B = vFloorClipPlane.y
C = vFloorClipPlane.z
D = vFloorClipPlane.w

The equation is normalized such that D is the height of the camera from the floor in meters. Using this equation we can detect the floor plane or even stair plane at the same time. To calculate the distance between any joint and the floor, the joint coordinates and floor plane equation can be applied to the following Eq. (2).

$$D \text{ (distance - joint to floor)} = \frac{|Ax+By+Cz+D|}{\sqrt{(A^2+B^2+C^2)}} \quad (2)$$

where: x, y, z are the coordinates of the joint.

B. *Fall identification*

For human fall identification from other activities of daily life we also need the velocity of joint and the joint distance, since some of the daily activities such as falling on floor and lying on floor do share the same characteristics. In such cases the changes of distance from head to floor possess similar pattern except the time it takes. Therefore we need to calculate the velocity of the joints in order to distinguish such movements.

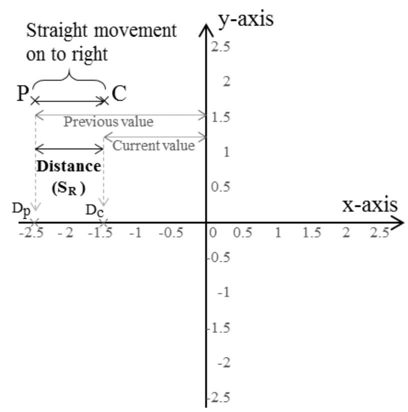
For velocity calculation, joint coordinates are extracted from every consecutive frames. The difference of distance from current frame and previous frame is calculated, which is than divided by the change of time as shown in Equation (3). The time difference is (1/30 seconds) between two frames since the sensor generates 30 frames per second. To compute the magnitude part for the velocity of right, left, up, down, coming close to sensor and going far to sensor movement, the distance is calculated by subtracting the current and the previous coordinates values on the respective axis, since these movements are just straight on to any of the axis as shown in Figure 1f, therefore the subtraction of current location from previous location will give the distance. For an example, Figure 1a shows the approach used for calculating the distance from a movement which is just straight to right side (SR) with respective to previous location. Here the distance is calculated by subtracting the current x-axis value from the previous x-axis value. Similarly any such movement to left side (SL shown in Figure 1(f) is also calculated by subtracting the current x-axis value from previous x-axis value. The other two straight movements (SU and SD), is calculated from the difference of current y-axis value from the previous y-axis value. SU calculation is illustrated in fig. 1b and SD is just the opposite movement to SU as shown in Figure 1(f).

$$\text{Velocity (V)} = \frac{D_c - D_p}{t_c - t_p} \text{ Meter/second} \quad (3)$$

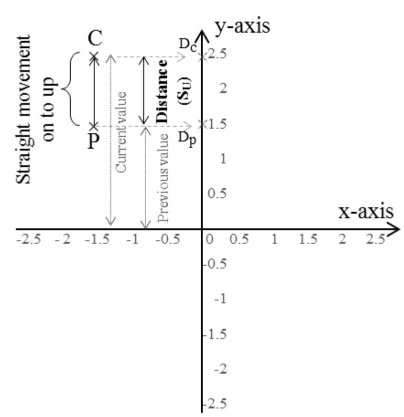
where: *Dc* is the Current Distance (current joint coordinate), *Dp* is the Previous Distance (previous joint coordinate), *tc* is the current time in second and *tp* is the previous time in second.

$$D = \sqrt{((y - y`)^2 + (x - x`)^2)} \quad (4)$$

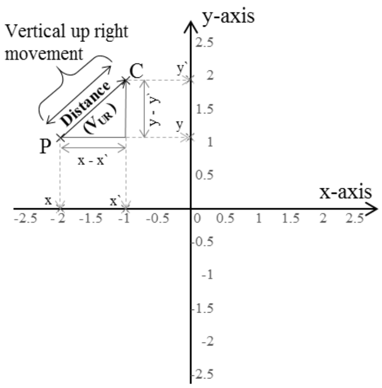
The distance for the other vertical movements are calculated using the formula shown in Equation (4), according to the description in Figure 1(d). Here we use the simple trigonometry formula: square of hypotenuse (d) is equal to square of opposite plus the square of adjacent. Opposite and adjacent is calculate using the current and the previous coordinates of x-axis and y-axis respectively. An example is shown in Figure 1(c), where the distance is from a movement which is vertically up to right side. For the other three vertical movements, the distance is calculated using the same approach as in Figure 1(c) and the distance (d) is calculated using the Equation (4). Once the distance is calculated, Equation (3), is used to find the velocity.



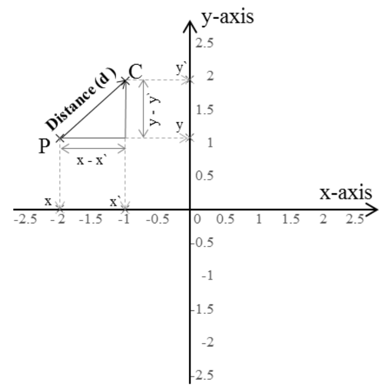
(a) Straight movements (left or right)



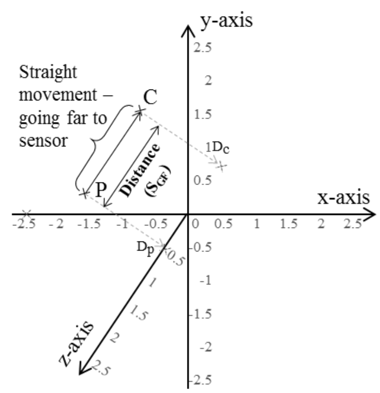
(b) Straight movements (up or down)



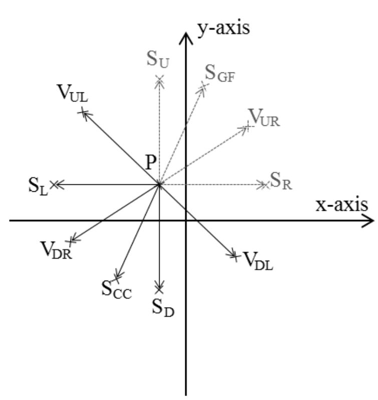
(c) Vertical movements



(d) General for Vertical movements



(e) Straight movements (coming close or going far)



(f) Other straight and vertical movements

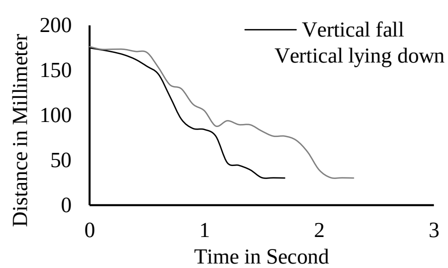
Figure 1: Coordinate system and velocity calculation description

The direction is obtained using the coordinate system of the Kinect. As per the coordinate system shown in Figure 1f, any movement to the right or top or going far (to any axis) gives a positive value for the distance difference. Similarly any movement to left or down or coming close gives a negative value for the distance difference. Using this concept, the direction is determined to all the movements shown in Figure 1f.

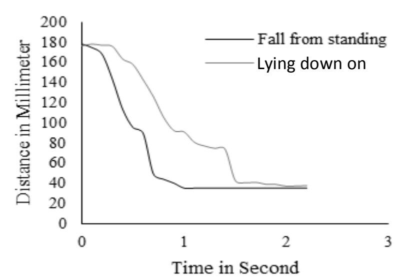
IV. Experimental Results

Our method has been tested on simulated falls and other Activities of daily life like walking, sitting on floor, bending, sitting on chair, falling from chair etc. From the trials conducted, our system was able to calculate the distance from head to floor and the velocity of head within the view of the sensor. Thus it was able to accurately distinguish falls from other activities using distance and velocities. Use of joint position together with the joint distance and velocity, improve the accuracy furthermore. Figure 2 shows the screen shot of activities corresponding to the results of the experimental activities illustrated in Figure 3.

The results shown in Figure 3, clearly distinguish falls (fall from standing and fall from chair) from other daily activities. This indicates that our system can differentiate falls from other activities even only using distance from head to floor. But this is true for the activities described in Figure 2 only. For activities like lying down on floor, which is very similar to a fall, distance from head to floor alone may not be able to separate the two activities. In such case we have to take into account the velocity of the movement to differentiate the two activities, since the changes in distance with respect to time has a clear gap in between a fall and lying on floor as seen in Figure 4.



(a) Falling and lying on floor vertically to sensor



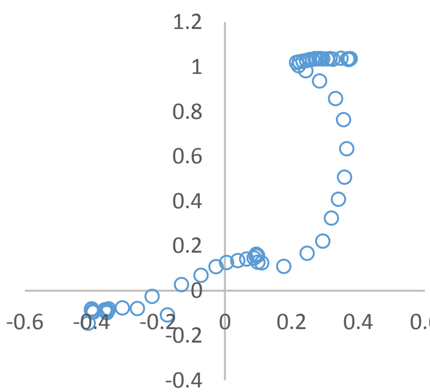
(b) Falling and lying on floor across the sensor

Figure. 4: Distance changes between a fall and lying on floor from standing

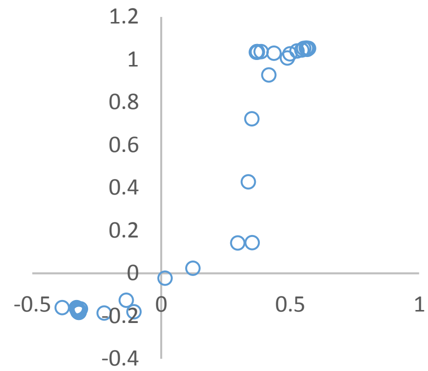
As seen in Figure 4, the distance changes for lying down on floor from standing has a very similar pattern to fall from standing except for the changes of distance over time. Therefore the velocity can be used to distinguish the two movements. Figure 5 shows the variation of velocity for a fall from standing and lying on floor from standing. It is obvious that the velocity for "falling" is higher and the changes are rapid than for lying on floor. This changes can be used in conjunction with distance from head to floor to separate the two activity.

Similarly, brutal movements do have analogous pattern with the same normal movement as far as the changes of distance from head to floor is concerned. An example of such a movement is shown in Figure 6, where the changes of distance for sitting brutally on floor is illustrated with a normal sitting on floor. Here the change of distance for the brutal movement is faster about half a second while sitting normally takes one and half second. Thus the change in distance over time is similar to a fall and an intentional lying on floor, therefore the velocity can be used to distinguish the brutal movement from normal sitting on floor.

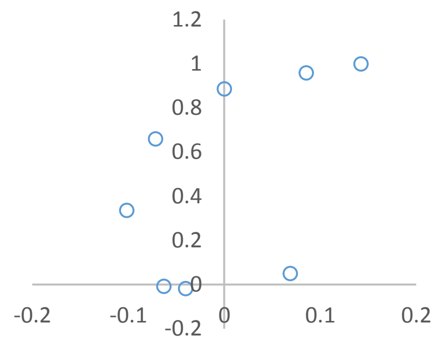
Apart from the distance and changes of velocity during any of the activities, the change of head position also gives important information in distinguishing fall from other activities. These changes in head position are taken into account while computing the joint position for confirming a detected fall movement from the changes of distance and joint velocity. The below Figure 7(a) and 7(b) illustrates the changes of head position during a fall from standing and lying down on floor from standing. Likewise, Figure (c) to (e) show the changes of head position from standing to brutally sit on floor, standing to a fall while trying to sit on floor and from standing to sit on floor respectively.



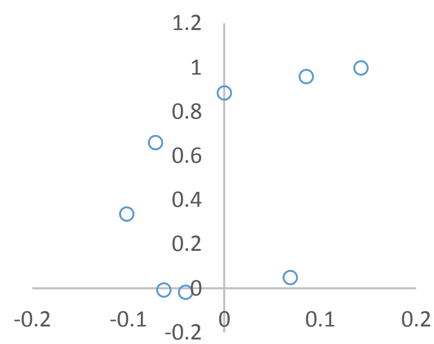
(a) During lying down from standing



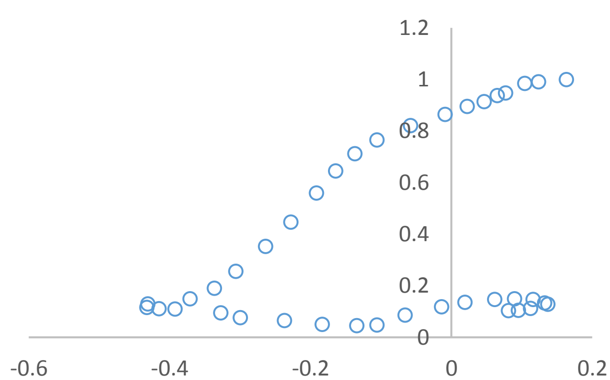
(b) During a fall from standing



(c) Sitting brutally on floor



(d) Fall while sitting



(e) Sitting on floor

Figure 7: Changes of head position

V. CONCLUSION

In this paper we proposed a human fall detection system based on human joint measurements using depth images generated by the Kinect infrared sensor. The experimental results showed that the algorithm used on the system can accurately distinguish fall movements from other daily activities with an average accuracy of 96.55%. The system was also able to gain a sensitivity of 100% with a specificity of 95%. The proposed system was able to distinguish all fall movements from other activities of daily life accurately. The velocity of joints greatly help to classify certain movements were the distance changes possess similar variation. The proposed system could be further improved by focusing on the algorithms for human extraction.

ACKNOWLEDGMENT

The Author Yoosuf Nizam would like to thank the Universiti Tun Hussein Malaysia for providing lab components and GPPS sponsor for his studies.

REFERENCES

[1] Griffiths, C., Rooney, C.,Brock, "A.: Leading causes of death in England and Wales–how should we group causes," *Health Statistics Quarterly*, vol. 28, no. 9, 2005.

[2] Baker, S.P.,Harvey, "A.: Fall injuries in the elderly," *Clinics in geriatric medicine*, (1985) 1(3): 501-512.

[3] Mohd, M.N.H., et al. "Internal state measurement from facial stereo thermal and visible sensors through SVM classification," *ARPN Journal of Engineering and Applied Sciences*, vol. 10, no. 18, pp. 8363-8371, 2015.

[4] Mohd, M.N.H., et al. "A Non-invasive Facial Visual-Infrared Stereo Vision Based Measurement as an Alternative for Physiological Measurement," *in Computer Vision-ACCV 2014 Workshops*, Springer, 2014.

[5] Bian, Z.-P., Chau, L.-P.,Magnenat-Thalmann, N. "A depth video approach for fall detection based on human joints height and falling velocity," *International Coference on Computer Animation and Social Agents*, 2012.

[6] Kepski, M.,Kwolek, B. "Fall detection using ceiling-mounted 3d depth camera," *in Int. Conf. VISAPP*, pp.2, 2014.

[7] Kwolek, B.,Kepski, M. "Human fall detection on embedded platform using depth maps and wireless accelerometer," *Computer methods and programs in biomedicine*, vol. 117, no. 3, pp. 489-501, 2014.

[8] Kwolek, B.,Kepski, M. "Improving fall detection by the use of depth sensor and accelerometer," *Neurocomputing*, vol. 168, pp. 637-645, 2015.

[9] Yang, L., et al.: New fast fall detection method based on spatio-temporal context tracking of head by using depth images," *Sensor*s, vol. 15, no. 9, pp. 23004-23019, 2015.

[10] Kawatsu, C., Li, J.,Chung, C. "Development of a fall detection system with Microsoft Kinect," *in Robot Intelligence Technology and Applications*, Springer, pp. 623-630, 2013.

[11] Gasparrini, S., et al. "A depth-based fall detection system using a Kinect® sensor," *Sensors*, vol. 14, no. 2, pp. 2756-2775, 2014.

[12] Mundher, Z.A.,Zhong, J. "A Real-Time Fall Detection System in Elderly Care Using Mobile Robot and Kinect Sensor,".

[13] Rougier, C., et al. "Fall detection from human shape and motion history using video surveillance," *in Advanced Information Networking and Applications Workshops*, *AINAW'07*, 2007.

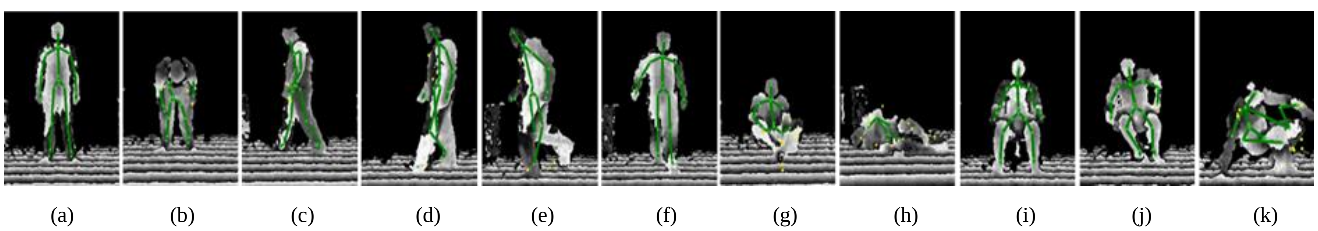


Figure 2: Screenshot of activities performed with corresponding depth images. (a) Standing; (b) bend and stand up; (c) Walking across the sensor; (d) Walking slowly across the sensor; (e) Running; (f) Walking around the sensor; (g) sit down on floor; (h) falling from standing; (i) sit on chair; (j) fall from chair; (k) stand up

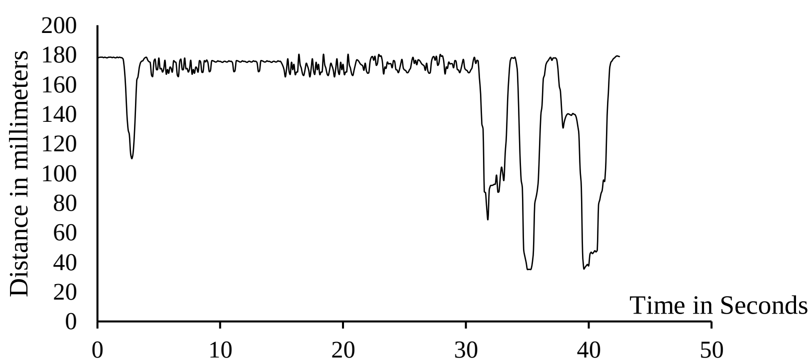


Figure 3: Result of changes in distance for the activities in above figure.

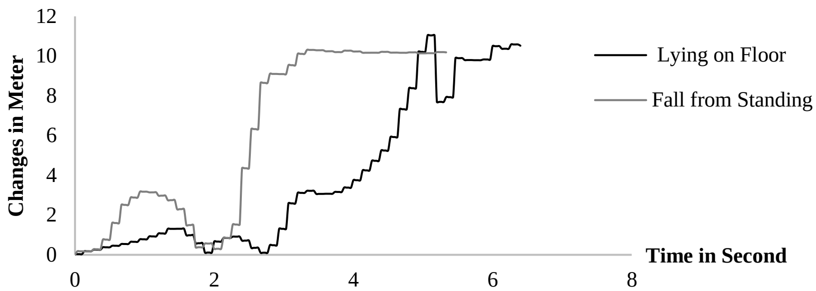


Figure 5: Changes of velocity for falling and lying on to floor

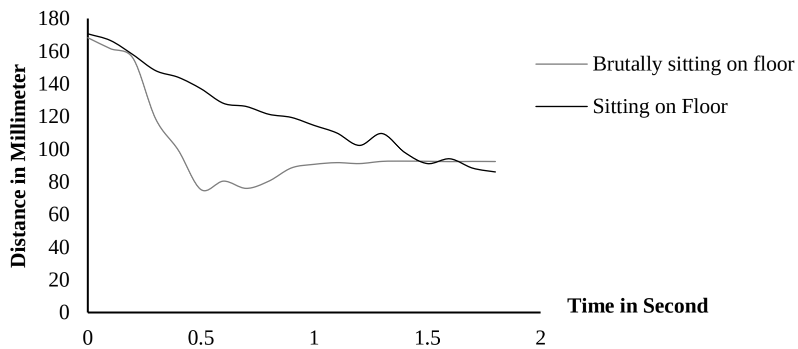


Figure 6: Changes of distance from head to floor for brutally and usual sitting on floor