Depth Sensor Based Skeletal Tracking Evaluation for Fall Detection Systems

Subarna Sinha[1], Suman Deb[2]
[1](Department of Computer Science And Engineering, NIT Agartala, India)
[2](Department of Computer Science And Engineering, NIT Agartala, India)

ABSTRACT ***:*** *Falls are very common in elderly due to various physical constraints. Since falls may cause serious injury and even death, fall detection systems are very important, especially when the victim is alone at home or is unable to seek regular/timely medical assistance. In this paper, development of a fall detection system based on Kinect sensor is evaluated. Microsoft Kinect is a low cost RGB-D sensor and it has the ability to track joint positions which could prove useful as a sophisticated tool for fall detection. In this study the potential of Kinect for application in fall detection has been investigated.*

Keywords *- fall detection, Kinect, skeleton detection, depth image, infrared camera.*

1. INTRODUCTION

Aging has always been a major social challenge in the entire world. Elderly people living alone in an apartment are very common these days. The reasons are manifold like social, economical, personal etc. The number of such elderly is large in number. They require assistive support in daily work. Involuntary falls are frequent for the elderly. Falls cause a loss in quality of life for the fallen elderly and can be more dangerous due to the fact that the victim can easily lose consciousness and thus become unable to seek help if they are home alone, which is detrimental to their long-term health if the accident is serious and undetected [1]. In order to avoid this, fall detection systems are required to be developed.

MICROSOFT KINECT is a RGB-D sensor developed initially by Microsoft for the Xbox game console [2]. It has the advantage of 3D motion capture algorithm which enables the user to play without even using a controller. It was launched by Microsoft in 2010. Within a few months of its launch many interesting applications using Kinect came into being. Researchers soon found out the extended capabilities of Kinect which included the depth sensing technology of Kinect to be used far beyond gaming. Moreover, being a low cost product compared to the traditional 3-D cameras like stereo cameras [3] and timeof- flight (TOF) cameras [4] Microsoft Kinect has become a widespread success. The success of Kinect in gaming field is solely because of its powerful sensors and robust skeletal tracker. The camera is co-located with the display as well which helps avoiding the range limitation of the depth sensor.

Available Fall detection systems have fall detection rates of 70-80% [5], which might be error prone and unconvincing to some consumers when making decisions on purchasing such systems. Other technologies such as the uses of floor sensors [6] have shown fall detection rates of about 90%, but may suffer from other problems that limit its use such as vulnerability to spillages, etc. We aim to evaluate the capability of Microsoft Kinect sensor for using in a fall detection system which determines automatically if a fall has occurred.

2. RELATED WORK

The Kinect is a small and affordable enough sensor to be used in any home environment, and it does not require patients to wear anything that could limit their movement. Other studies have also identified the Kinect's potential for use in physical therapy.

A lot of research work has been done for developing fall detection systems in recent years[7]. However, despite many efforts undertaken to achieve reliable fall detection [8], the existing technology does not meet the seniors' needs [9]. Particularly, the available solutions generate too much false alarms. These false alarms can be annoying to anyone, not only the elderly. Most of the currently available techniques for fall detection are based on body-worn or built-in devices. They typically employ accelerometers or both accelerometers and gyroscopes [8]. Such sensors may generate false alarms getting confused between fall and fall like activities [10]. Moreover, the detectors that are typically worn on a belt around the hip, are uncomfortable as well especially during the sleep[11]. The wide range of research works in this field include pressure pads [12], single CCD camera [13], multiple cameras [14], specialized omni-directional ones [15] and stereo-pair cameras.

3. KINECT AND ITS FEATURES

Kinect software deals with the tools available for working with Kinect hardware as well as various libraries to use its functionalities for developing several applications. Several tools for Kinect like- Microsoft Kinect SDK [16], OpenNI [17] and OpenKinect [18] etc. are available for development of applications. OpenNI 2.0 is the highest version of OpenNI SDK. Microsoft Kinect SDK, released by Microsoft and its current version is 1.8. OpenKinect is a free, open source library maintained by an open community of Kinect people. The Microsoft SDK is only for working in Windows platform but OpenNI is a multiplatform and an open-source tool.

Fig. 1. Kinect Sensor

Inside Kinect sensor device we have the following hardware components:

3.1. IR emitter and IR depth sensor: An infrared (IR) emitter and an IR depth sensor.IR beams are emitted by the emitter. These beams when reflected back to the sensor are read by the depth. The reflected beams are converted into depth information measuring the distance between an object and the sensor. This makes capturing a depth image possible. The depth sensor contains a monochrome CMOS sensor and infrared projector that help create the 3D image of the environment.

3.2. Color sensor: It is an RGB camera which captures RGB pixel information of the facial and body information. Its pixel resolution is 640x490 and has a frame rate of 30fps.

3.3. Tilt motor: This motor is pivot of the device using this motor we can move Kinect sensor device 27degree up and down.

3.4. Microphone array: This microphone array consists of 4 microphone used to record audio as well as to find the audio location.

Kinect camera uses a diffractive optical element and an infrared laser diode to generate an irregular pattern of dots basically the same pattern is repeated 3x3 times. It includes a color and a two megapixel grayscale chip with an IR filter, which is used to determine the disparities between the emitted light dots and their observed position. In order to triangulate the depth of an object in the scene the identity of an observed dot on the object must be known. The irregular pattern can be generated with much more certainty than with a uniform pattern. The camera uses a 6mm lens and produces a 640x490 pixel sized raw depth map which consists of an 11-bit integer value per pixel. The depth values describe the distance to the imaginary image plane and not to the focal point.

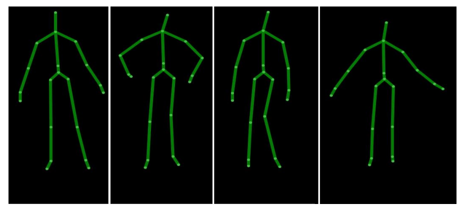
Fig. 2. Skeletal tracking using Kinect for Windws toolkit.

Human as well as Hand gesture recognition have been a matter of discussion and interest among researchers in the field of computer vision. With the invention of Kinect the research has taken a turn towards the use of RGB-D data provided by Kinect. Instead of extracting local features from raw video sequences, the new algorithms are provided with the RGB and depth images and as well as the skeletal joints. The research areas involved skeletal tracking or pose estimation and activity recognition etc. Kinect has the feature of being an advanced skeletal tracker, which ignites the opportunities in the field of human activity analysis. A core part of the algorithm is described in the paper[19] in which they estimate the body joint positions by calculating the local centroids of the body part probability mass. The direct correspondences between image pixels and a 3-D mesh model can be found in some works[20] where energy minimization is used for pose detection.

4. SKELETAL TRACKING AND FALL DETECTION

The Kinect sensor consists of an infrared laser emitter, an infrared camera and an RGB camera. The inventors describe the measurement of depth as a triangulation process [21]. The laser source emits a single beam which is split into multiple beams by a diffraction grating to create a constant pattern of speckles projected onto the scene. This pattern is captured by the infrared camera and is correlated against a reference pattern. The reference pattern is obtained by capturing a plane at a known distance from the sensor, and is stored in the memory of the sensor. When a speckle is projected on an object whose distance to the sensor is smaller or larger than that of the reference plane the position of the speckle in the infrared image will be shifted in the direction of the baseline between the laser projector and the perspective center of the infrared camera. These shifts are measured for all speckles by a simple image correlation procedure, which yields a disparity image. The Microsoft SDK supports tracking of 20 joints, shown in Figure 3. It provides X, Y, and Z-coordinates in meters from the sensor for each joint position, according to the axes as shown in Figure 4. The labels in the figure are placed on the positive direction of each axis.

Microsoft's joint tracking algorithm identifies joint positions by processing a depth image. The algorithm first guesses a particular joint position for a pixel in the depth image. Also a confidence level for each pixel is calculated. After this, it chooses the skeleton that is most likely given those joint labels and confidence levels. But before this could be done, the algorithm had to know how to make accurate guesses for joint positions.

The implementation of this project had two main purposes. First was to demonstrate that it would be possible to develop useful fall detection software tools with the Kinect. Another was to collect data for analysis. The programs written are not in a form that could be useful for fall detection. However, further work could easily be done to improve on these proof-of-concept applications to enable their use to facilitate the patients or elderly.

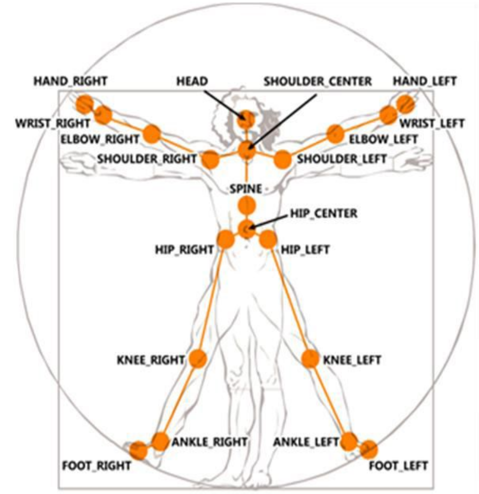

Fig.3. The skeleton joints positions in Microsoft Kinect SDK [22]

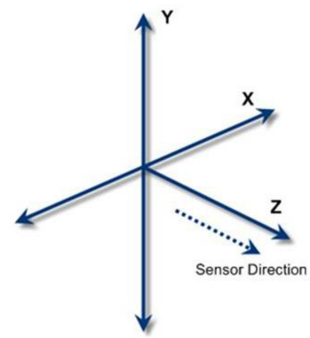

Fig. 4. The coordinate system in Microsoft Kinect SDK [22]

5. RESULTS

All testing was performed using a HP laptop running Windows 7 with a core i3 Intel processor and 3 GB of memory. Most of the experiments were conducted with the test subject at a distance of about 2 meters from the Kinect sensor. Also testing has been done to check the effect of distance variation. In most cases the subject directly faced the Kinect sensor, beside which the laptop screen displayed the running program with movements and rendered skeleton. A person sitting next to the laptop screen (outside the view of the Kinect) gave instructions to the subject. The top view of the lab layout is given in figure below.

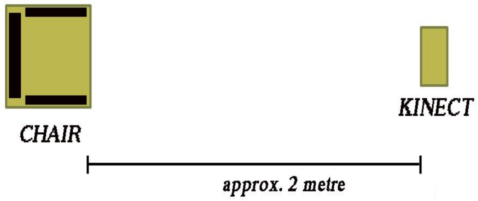

Fig. 5. Overhead View of Lab

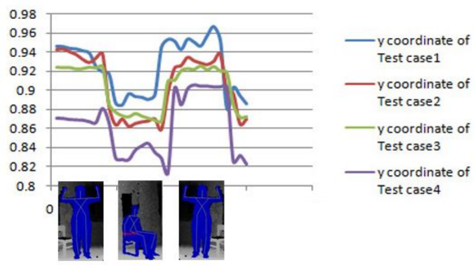

Fig. 6. Head Position and Regions of Movement

The test was performed on four test cases as shown in the figure 6. The y coordinates of the skeletal data obtained from Kinect have been used. The test cases were performed on four different people (volunteers) of different size and height. The graph above has been plotted by taking y coordinates of the head joint position on the y axis and frame numbers on the x axis. Of all the joints tracked by Kinect, we found that tracking head joint is relevant for use in fall detection problem. This is because of the natural method of fall of a person

The test was performed as follows: The subject started out sitting on a chair facing the Kinect sensor. About after four seconds of recording, the volunteer stood up and remained standing still for four more seconds. For each of the test cases, joint positions were recorded for all the skeletal joints and the instantaneous frame rate. This data was written to a text file for later analysis. One area of analysis in these experiments explored the feasibility of identifying phases of movement in fall detection from the program's output data. Given a set of points representing the position of a patient's head during a fall, it was clearly identifiable when the patient was sitting, or standing.

The data was collected using the Microsoft Kinect SDK for Windows, the corresponding activity at a particular frame has been shown with the corresponding posture on the graph. It is interesting to see the dip in the graph where the subject sat down. The sitting, sitting-to-standing, and standing regions are also clearly distinguishable from the graph. And since all the four test cases had the same nature of graph, the deduction is quite evident.When the subject was too close or too far away from the sensor, the inconsistency of the joint positions made the phases of movement somewhat distorted, but even at these distances it was still not excessively difficult to make an approximate estimation. The ability to separate regions of movement was very important for further analysis of the data in this study, and it is also crucial for the successful use of the Kinect as a fall detection tool. The results of this section of analysis suggest the Kinect's output data is well-suited for this purpose.

6. CONCLUSION

In this investigation of the potential of the Microsoft Kinect as a tool for fall detection, it was found that the data and capabilities of the device are very promising, particularly when only one person is in the field of view. Joint position data was gathered from four subjects of varying genders, heights, and body types using the developed application while the subjects performed few positional changes (e.g. sitting, standing, moving, etc) in front of Kinect. The results were easily identifiable from graphs of position data from the Kinect. Although there are many possibilities of refinements and enhancements that could be done to the existing prototype applications to make them more suitable for use in a clinical setting, the application could successfully determine that the Kinect has much potential for use in fall detection as a tool for patients or elderly.

REFERENCES

[1] T. Dutta, Evaluation of the kinect sensor for 3-d kinematic measurement in the workplace ,e. *Applied Ergonomics*, October 2011.

[2] T. Stoyanov, A. Louloudi, H. Andreasson, A. J. Lilienthal. *Comparative evaluation of range sensor accuracy in indoor environments*, Evaluation, 2011.

[3] S. Gokturk, H. Yalcin, and C. Bamji, A time-of-flight depth sensorsystem description, issues and solutions, in *Proc. IEEE Conf. Comput.Vision Pattern Recognit*. Workshops, 2004, pp. 35–45.

[4] Perry, J. T., Kellog, S., Vaidya, S. M., Youn, J. H., Ali, H., & Sharif, H. (2009, December). Survey and evaluation of real-time fall detection approaches. In *High-Capacity Optical Networks and Enabling Technologies (HONET)*, 2009 6th International Symposium on (pp. 158-164). IEEE.

[5] Rimminen, H., Lindstrom, J., Linnavuo, M., & Sepponen, R. (2010). Detection of falls among the elderly by a floor sensor using the electric near field. *Information Technology in Biomedicine, IEEE Transactions* on, 14(6), 1475-1476.

[6] Mubashir, M., Shao, L., Seed, L.: A survey on fall detection: Principles and ap- proaches. *Neurocomputing 100*, 144 { 152 (2013), special issue: Behaviours in video

[7] Noury, N., Fleury, A., Rumeau, P., Bourke, A., ‼ OLaighin, G., Rialle, V., Lundy, J.: Fall detection - principles and methods. In: *Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society*. pp. 1663{1666 (2007)

[8] Yu, X.: Approaches and principles of fall detection for elderly and patient. In: 10th *Int. Conf. on e-health Networking, Applications and Services*. pp. 42{47 (2008)

[9] Bourke, A., O'Brien, J., Lyons, G.: Evaluation of a threshold-based tri-axial accelerometer fall detection algorithm. *Gait & Posture 26(2),* 194{199 (2007)

[10] Degen, T., Jaeckel, H., Rufer, M., Wyss, S.: Speedy: A fall detector in a wrist watch. In: *Proc. of IEEE Int. Symp. on Wearable Computers*. pp. 184{187 (2003)

[11] Tzeng, H.W., Chen, M.Y., Chen, J.Y.: Design of fall detection system with oor pressure and infrared image. In: Int. Conf. on System Science and Engineering. pp. 131{135 (2010)

[12] Anderson, D., Keller, J., Skubic, M., Chen, X., He, Z.: Recognizing falls from silhouettes. In: *Annual Int. Conf. of the Engineering in Medicine and Biology Society*. pp. 6388{6391 (2006)

[13] Cucchiara, R., Prati, A., Vezzani, R.: A multi-camera vision system for fall detection and alarm generation. *Expert Systems 24(5)*, 334{345 (2007)

[14] Miaou, S.G., Sung, P.H., Huang, C.Y.: A customized human fall detection system using omni-camera images and personal information. *Distributed Diagnosis and Home Healthcare* pp. 39{42 (2006)

[15] Jansen, B., Deklerck, R.: Context aware inactivity recognition for visual fall detection. In: *Proc. IEEE Pervasive Health Conference and Workshops*. pp. 1{4 (2006)

[16] Microsoft Kinect SDK [Online]. Available: http://www.microsoft.com/en-us/kinectforwindows/.

[17] OpenNI [Online]. Available: http://www.openni.org/

[18] OpenKinect[Online].Available:https://github.com/OpenKinect/libfreenect/.

[19] J. Taylor, J. Shotton, T. Sharp, and A. Fitzgibbon, “The vitruvian manifold: Inferring dense correspondences for one-shot human pose estimation,” in *Proc. CVPR, 2012*, pp. 103–110.

[20] J. Shotton, A. Fitzgibbon, M. Cook, T. Sharp, M. Finocchio, R. Moore, A. Kipman, and A. Blake, “Real-time human pose recognition in parts from a single depth image,” in *Proc. IEEE Conf. Comput. Vision Pattern Recognit., 2011*, pp. 1297–1304.

[21] S. Rusinkiewicz, M. Levoy,” Efficient variants of the ICP algorithm”. In *Third International Conference on 3D Digital Imaging and Modeling*, 2001.

[22] Microsoft. (2011, July) Kinect for Windows SDK Beta 1 Programming Guide Version 1.1. pdf.