INTEGRATING SMARTPHONE AND KINECT FOR FALL DETECTION

Hone-Jay Chu[1], Guang-Je Tsai[2]

Department of Geomatics Engineering, National Cheng Kung University, No. 1, Daxue Road, East District, Tainan 701, Taiwan

[1]honejaychu@gmail.com, [2]tpp1114@gmail.com

Abstract

Smartphones are widely applied in various applications because of the recent developments in mobile sensing and wireless communication technologies. For senior care, location-based fall detection and alarm systems are necessary. This study applies a smartphone as a location-based platform and considers accelerometers to detect the fall in the aged. On the basis of three-axis accelerometers, one can determine the state of a fall. Moreover, Kinect is used to assist the smartphone fall detection system and to reduce false detections in the smartphone fall detection system. With depth data collection, this study extracts the skeletal joints from the color and depth streams for fall detection. Kinect and smartphone sensors can be combined to obtain sufficient information. When a fall occurs, the magnitude suddenly increases to more than 2.5 g and the tilt angle variation exceeds 40°. The Kinect sensor is provided the skeleton model and the images to recognize the fall for double-check . The images obtained during the falling motion can be transmitted to the web server. Then, the proposed server can indicate the location of the detected fall. Activity monitors are used to monitor fall movements from the web server.

Keywords: smartphone, Kinect, fall detection

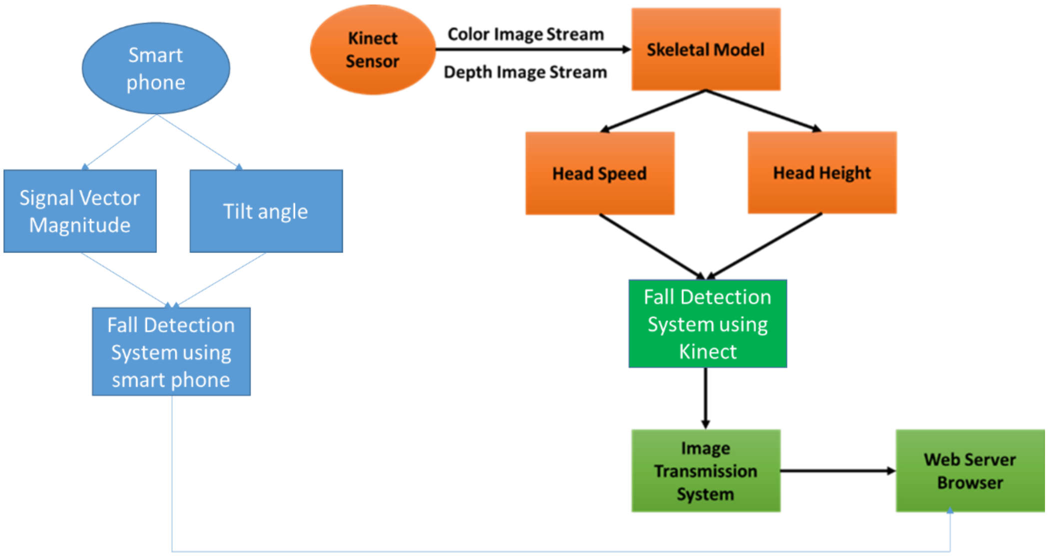
1. INTRODUCTION

Fall detection systems are divided into three main categories, namely, wearable, ambience-based, and vision-based devices (Mubashir et al., 2013; Igual et al., 2013). Posture, inactivity, and motion can be detected by these devices. The number of systems for detecting these aspects of fall has increased dramatically. Smartphones and many kinds of sensors, such as global positioning satellite (GPS) sensors, magnetic sensors, and accelerometers, have been used widely in recent years (Miao et al., 2015). As a smartphone sensor, the accelerometer can be used to determine and monitor the falling motion efficiently. The smartphone is designed to require low battery power consumption. Figueiredo et al. (2016) explored the fall information provided by the accelerometer, magnetometer, and gyroscope sensors in a smartphone. A simple and reliable algorithm for fall detection is proposed using a threshold-based approach. Aguiar et al. (2014) presented a smartphone-based fall detection system that uses a combination of acceleration information derived from machine learning classification. Abbate et al. (2012) recognized fall-like activities of daily living on the basis of the acceleration patterns of smartphones and significantly reduced the number of false alarms.

Kinect is a low-cost sensor in which the device features a red, green, and blue (RGB) color camera, a depth sensor using an infrared (IR) projector, and an IR camera, which can also help detect the location of a person in three-dimensional (3D) space (Han et al., 2013). Gasparrini et al. (2014) proposed an automatic fall detection method based on the Kinect sensor. The tracking algorithm allows one to extract the fall motion of a human subject. Yang et al. (2016) presented a fall detection method of elderly people in a room on the basis of shape analysis of 3D depth images captured by a Kinect sensor. Mastorakis and Makris (2014) measured the falling velocity based on the 3D bounding box. The system can detect falls accurately on walk. Our study proposed the skeleton-based fall detection method using the Kinect sensor. The mere use of portable sensors may not be sufficient to detect accidents immediately and to detect the falling motion of older adults. In recent years, a depth sensor has been widely developed, and the cost is rapidly reduced. The depth sensor includes two cameras for capturing depth information and image data. Stone and Skubic (2015) proposed an approach that adopts the Kinect sensor with decision tree for fall detection. However, depth information is insufficient for accurately detecting fall incidents.

This study considers the skeletal model from Kinect SDK and the accelerometer signal from smartphones to improve the accuracy of fall detection. The flowchart is shown in Figure 1. In this study, the combination of smartphone and Kinect sensors can provide fall detection information. These fall detection systems depend on motion information and fall characteristics, such as tilt angle (TA) and signal vector magnitude (SVM) from smartphones (He et al., 2012). Moreover, the height and velocity of the head can be used to detect the falling motion of a user through Kinect. If the head height is lower than the threshold, then the fall can be detected. At the same time, the images of the fall can be transmitted to the web server. If one man falls down, then the GPS location and corresponding images are shown from the web server. The web server sends a notice to the family and health center immediately. Through this process, users can monitor the fall of elderly people using the web server.

Figure 1: Flowchart of fall detection



2. MATERIAL AND METHOD

2.1. Microsoft Kinect sensor

Microsoft's Kinect sensor was released five years ago. With the novel technology in depth data collection, Kinect provides researchers one aspect lacking in traditional studies. The Kinect sensor is originally designed for the gaming console Xbox 360, which is completely different from traditional consoles. The sensor is characterized by its ability to acquire depth data. Depth data can be transformed into a color point cloud for many applications. The Kinect sensor is a composite device with several sensors (Figure 2), such as color (RGB) camera, IR emitter, and IR camera. The Kinect can obtain RGB and IR images at 640 × 480 pixel resolution and 30 frames per second (FPS) (Figure 3). The depth image also provides practical information for fall detection. The specifications of Kinect are shown in Table 1.

Figure 2: Microsoft Kinect sensor

Figure 3: RGB (left), IR (middle), and depth (right) images obtained by Kinect

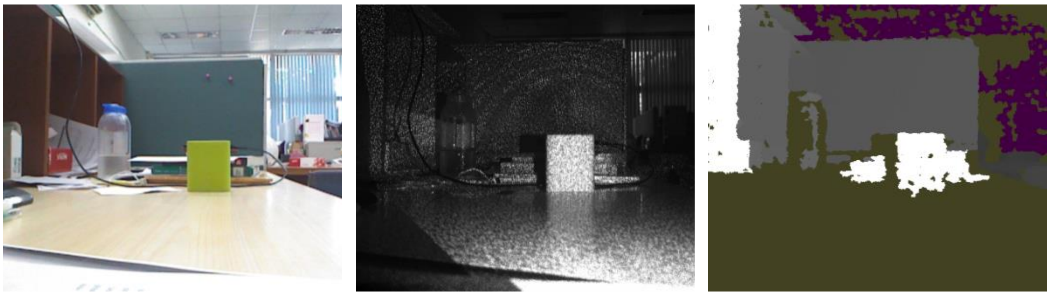


Table 1: Kinect specifications

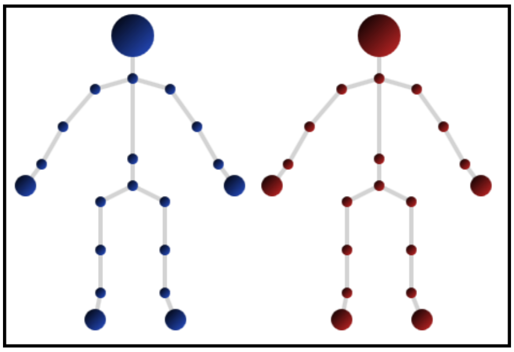
Viewing angle (VFOV/HFOV)	43° vertically/57° horizontally
Vertical tilt range	±27°
Frame rate (depth and color stream)	30 FPS
Color image resolution	1,280 × 960 pixels at 12 FPS 640 × 480 pixels at 30 FPS
Depth image resolution	640 × 480 pixels at 30 FPS 320 × 240 pixels at 30 FPS 80 × 60 pixels at 0 FPS
Depth sensor range	1.2 m to 3.5 m

2.2. Skeletal Model

With the rapid advancement of computer vision technology, many studies have focused on human motion recognition (Aarai and Andrie, 2013). The complexities of human motion recognition have increased because of high-dimensional data and intra-class variability caused by scale, viewpoint, and illumination (Wang et al., 2013).

On the basis of the images obtained from the IR and RGB cameras, this study extracts the skeletal joints from color and depth streams. The depth stream includes the data on coordinates in each frame and incorporates these data with the color stream to detect the human skeleton joints by computer vision methods. The skeletal model of body parts are defined to be localized near 20 skeletal joints in Figure 4. Thus, the 3D locations of the skeletal joints can be applied for fall detection.

Figure 4: Skeleton tracked by Kinect

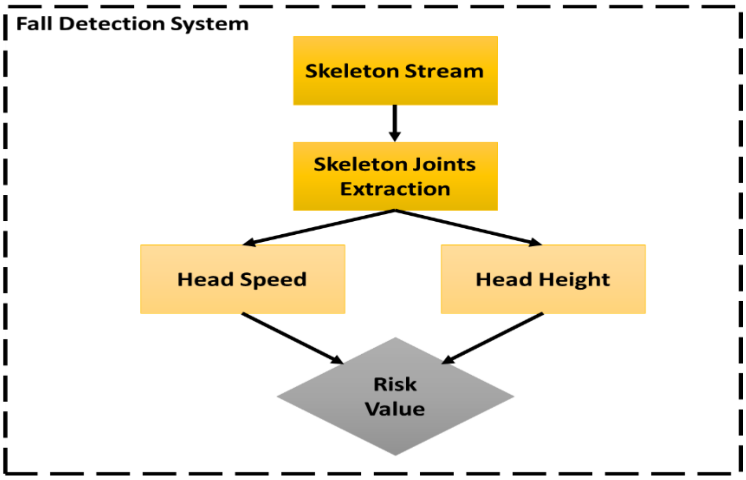


2.3. Skeleton-based Fall Detection System

Figure 5 shows the skeleton model in the fall detection system. The skeleton model is divided into frames corresponding to the depth and color streams to analyze the skeleton joints in sequential time. After requiring a skeleton stream from the skeleton model, the points of interest (e.g., head) are extracted to calculate the risk value. In a fall detection system, two key parameters, such as head speed and head height, determine the risk value. Head speed is derived from the coordinates of head joints based on the displacement of the head joint within two frames. Meanwhile, head height is determined by the offset between the joints of the head and feet.

Compared with regular movements, a fall is assumed to occur with fast head speed (>0.2 m/s) and low head height (<50 cm). Actually, when a fall ensues, the risk value increases with time. If the risk value is more than a certain threshold, then the system recognizes the fall incident.

Figure 5: Fall detection by Kinect



2.4. Fall Detection System using a Smartphone

A simple formula can be used for fall detection using smartphone accelerometers. The coordinate system of the smartphone is shown in Figure 6. SVM can be represented as

$$SVM = \sqrt{(a_x)^2 + (a_y)^2 + (a_z)^2}, \qquad (1)$$

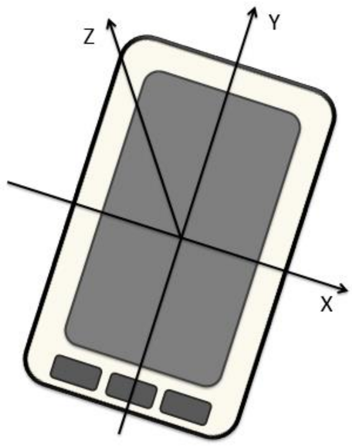
where a_x is the *x*-axis acceleration, a_y is the *y*-axis acceleration, and a_z is the *z*-axis acceleration. When the SVM value is greater than the threshold, a fall is assumed to occur. The test results identified SVM as the most suitable parameter at a threshold of 2.5 g (He et al., 2012).

TA was defined as

$$TA = \arcsin(a_y / \sqrt{(a_x)^2 + (a_y)^2 + (a_z)^2}). \qquad (2)$$

If TA variation is higher than the threshold of 40°, then it is classified as fall or lie.

Figure 6: Three smartphone axes



2.5. Image Transmission and Web Server Browser

Displaying real-time information is important in monitoring older adults in real time. Furthermore, the color stream is also helpful in acquiring the details of the fall. Figure 7 shows the Kinect GUI of the fall detection system. The upper-left panel presents the head speed of the user, the upper-middle panel shows the head height information, and the upper-right panel displays the risk value based on head speed and height. In the bottom portion of Figure 7, the color image is at the left side and the depth image with skeletal model is at the right side. The color stream and the depth stream with the skeletal model are also represented on the GUI of the fall detection system.

In this study, the web server is implemented to integrate two fall detection systems. The smartphone-based fall detection system provides the SVM and TA information to analyze the falls. For emergency demand, understanding the position of the scene is important for decision making. Therefore, a smartphone also sends the GPS data to the web server and records all sensor data. In the Kinect fall detection system, the system can store the sequential images. Once the accident occurred, the system sends the immediate images at the time of the accident to the web server. These images can help increase the understanding of such accidents and recover the details of the fall. Figure 8 presents the GUI of the web server browser. The left panel presents the detection information from the smartphone, and the right panel shows the sequential images from the Kinect fall detection system.

Figure 7: Kinect GUI for fall detection

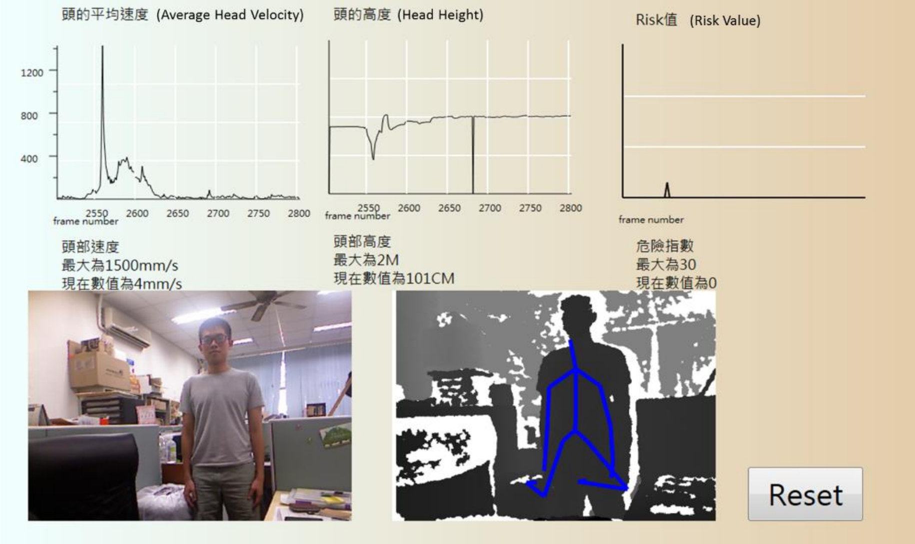
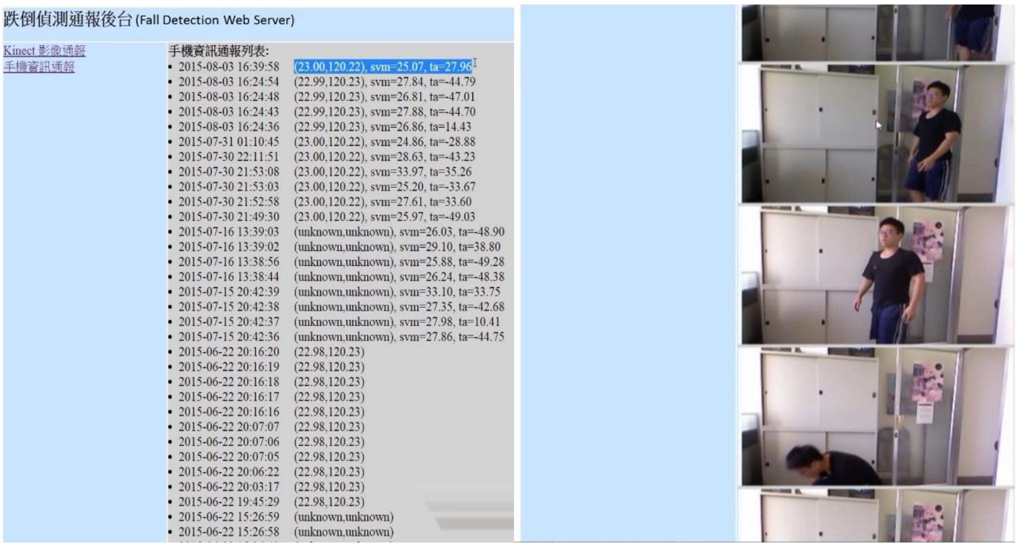


Figure 8: Web server GUI



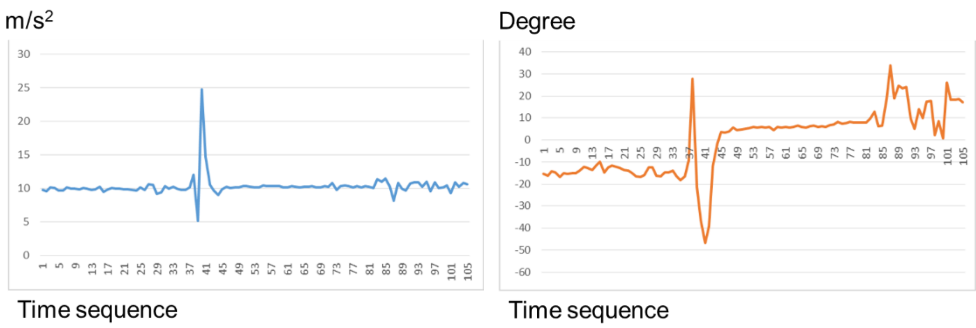
3. RESULTS AND DISCUSSION

In this study, two fall detection systems are integrated to analyze fall accidents transmitted to the web server. The experiments considered two highly confusing cases, falling and sitting, to evaluate the performance of the integrated fall detection system.

3.1. Case of Falling

In this case, the fall motion is detected by the detection system. The user holds the smartphone and simulates a fall. As mentioned in the section on the smartphone fall detection system, the SVM value reflects the acceleration magnitude of the fall. The TA of the device with respect to the user is shown. If the user falls down, then the SVM and TA values exceed the threshold. Figure 9 shows the SVM and TA data with the time recorded from smartphones. The *x*-axis is the time and the *y*-axis is the SVM and TA values. The plot shows that the SVM suddenly increases to 25 m/s^2 over 2.5 g and the TA variation exceeds 40° when the user falls. The sequential images are sent to the web server to display the incident.

Figure 9: Temporal SVM and TA (m/s^2 and degree, respectively) for falls



The Kinect detection system is used to assist the smartphone fall detection system to ensure that the fall detection system completely detects the accident. The Kinect sensor provides the skeleton model and the images to recognize the falling motion. Figure 10 presents the head velocity and height when the fall accident occurs. The figure shows the irregular fluctuation of head speed and the rapid descent of head height. In the system, the falling risk value increases linearly with time. When the risk value is higher than the threshold (e.g., 1 min), the line color is transformed from black to red to remind the user when the accident occurs. In Figure 11, all information is sent to the web server to evaluate the fall system. The image data from Kinect and falling time is sent to the web in the left and right panels. Through double recognition, the integrated system is more precise than the use of only one system for fall detection.

Figure 10: Head velocity and height in a fall case

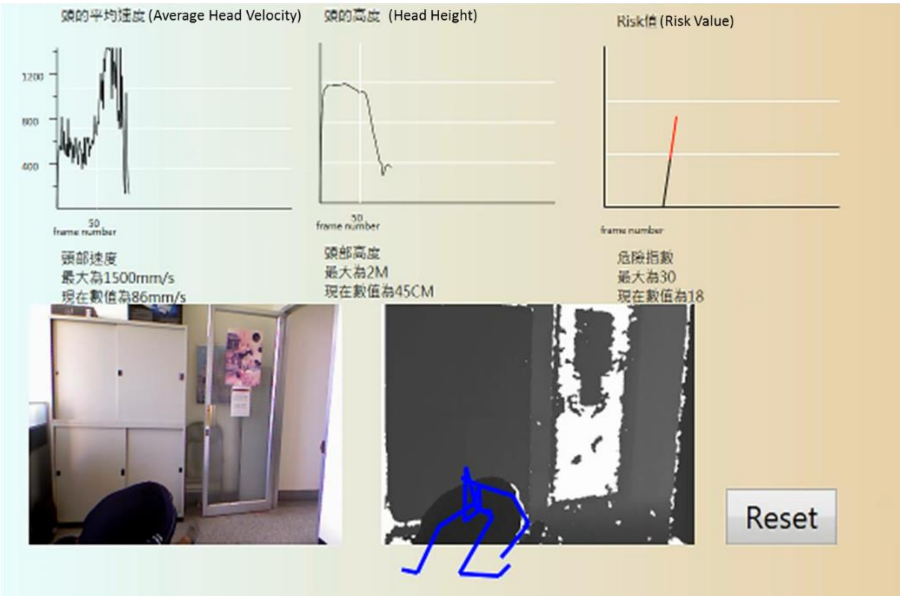


Figure 11: Web server in a fall case

3.2. Case of Sitting Down

In this study, the integrated system is simulated and tested in various cases. In the case of sitting down, the smartphone-based detection system is used to detect a fall. When the user suddenly sits down on a chair, a strong SVM value is detected from the smartphone. The TA value is also affected by a sitting down movement. The value change of the TA rapidly increases (Figure 12). However, if the Kinect is used to check for fall detection, then the system only recognizes the action as a simple sitting down movement. In Figure 13, the color image clearly shows the user sitting on a chair. Although the head speed is randomly increased, the stable head height renders the risk value lower than the threshold. Finally, the images from the web server are shown as cases of sitting down.

This study introduces a real-time algorithm that utilizes the human 3D skeleton joints expressed in world coordinates. Head height and velocity are acquired from the skeleton joint from Kinect. The information can be used to determine whether a particular activity is a fall or not. The fall detection system can accurately and robustly detect a fall when it occurs without false detections. False detection can be reduced by the combination of the smartphone and Kinect sensors. Only the smartphone detects falsely. If the two approaches are integrated, the system can view the actual state of the fall motion. After system detection, the web server displays the corresponding information on a family member or friend’s smartphone. The system provides the information of the location of the user and sends a notice to the family, friends, or help center of the subject immediately.

Figure 12: Temporal SVM and TA (m/s^2 and degree, respectively) for sitting down

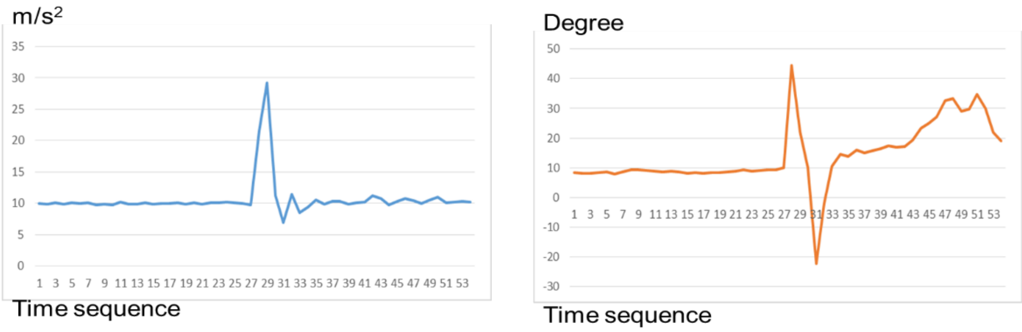
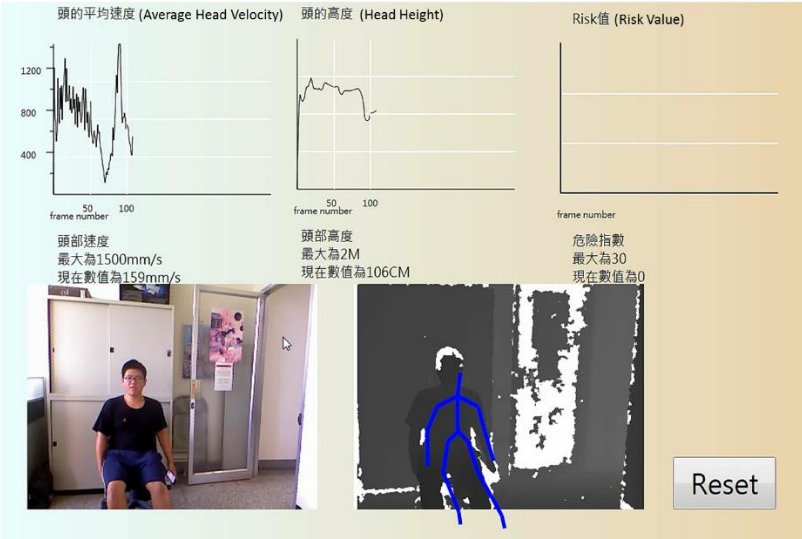


Figure 13: Head velocity and height for a sitting down case



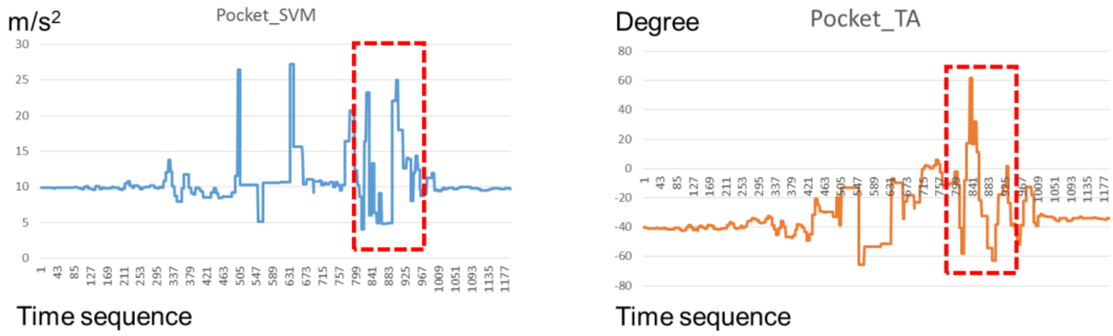
3.3. Discussion

Figure 14 shows the fall detection cases of the subjects with smartphones placed in their shirt or pant pockets. The test subject then walks and falls. When the fall occurs within the red window box, the SVM suddenly increases to 25 m/s^2 over 2.5 g and the TA variation exceeds 40°. The results show that such falls can be detected roubustly. A robust fall detection system must be highly reliable. The detection rate decreases relative to that in an experimental environment when the system is applied to a real situation. Thus, producing reliable results for actual applications becomes problematic (Igual et al., 2013; Feldwieser et al., 2014). Our study proposes a detection approach that combines information from smartphone and Kinect sensors, and the accuracy of the results is reinforced by visual verification from the server. However, sensor-based approaches are used in automatic fall detection systems (Mubashir et al., 2013). For example, using smartphone sensors is a pervasive and economical strategy. However, such approach is vulnerable to inconsistencies in various cases and is challenged in distinguishing between similar motions, such as falling and sitting. By using Kinect sensors, the algorithms can analyze head velocity and height. This method reduces the probability of false detections. In this study, dual models are applied and projected to minimize false judgments. The Kinect sensor is fixed in a room environment, whereas the smartphone is carried by the subject. Each of the systems contributes in achieving accurate detection. The Kinect sensor provides information to help the system diminish errors. Given the power requirement, the Kinect sensor can be fixed in a highly lit area.

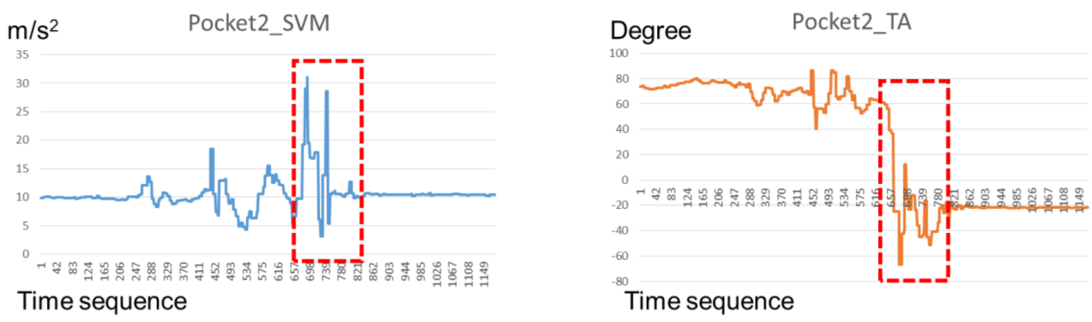
Multi-camera systems are advantageous because the cameras complement each other for to achieve robust detection (Sathyanarayana, 2015). Multiple Kinect sensors are projected to be used in fall detection in the future. With the increasing pervasiveness of microsensors and infrastructure, the detection method can be extended to high-resolution and real-time systems.

Figure 14: Fall detection by smartphones placed in the pockets of (a) pants and (b) shirts

(a)



(b)



4. CONCLUSION

This study integrates Kinect and smartphone to improve performance in fall detection. This presentation discusses the fall detection system and provides a prototype implementation to highlight the effectiveness of integrating sensor information for fall detection. Smartphone and Kinect sensors are combined to obtain sufficient fall information. Smartphone data are used to analyze and detect fall motion. The Kinect sensor provides the skeleton model and the images for recognizing falls. Through both means, the fall detection system can robustly detect a fall with a reduced rate of false detections.

5. ACKNOWLEDGMENT

The authors would like to thank the Ministry of Science and Technology of Taiwan for the financial support (103-2119-M-006-010).

6. REFERENCES

Aarai, K., Andrie, R. (2013). 3D Skeleton model derived from Kinect Depth Sensor Camera and its application to walking style quality evaluations. *International Journal of Advanced Research in Artificial Intelligence*, *2(7)*: 24-28.

Abbate, S., Avvenuti, M., Bonatesta, F., Cola, G., Corsini, P., & Vecchio, A. (2012). A smartphone-based fall detection system. *Pervasive and Mobile Computing*, *8*(6): 883-899.

Aguiar, B., Rocha, T., Silva, J., & Sousa, I. (2014). Accelerometer-based fall detection for smartphones. *Medical Measurements and Applications (MeMeA), 2014 IEEE International Symposium*. 1-6.

Feldwieser, F., Gietzelt, M., Goevercin, M., Marschollek, M., Meis, M., Winkelbach, S., ... & Steinhagen-Thiessen, E. (2014). Multimodal sensor-based fall detection within the domestic environment of elderly people. *Zeitschrift für Gerontologie und Geriatrie*, 47(8): 661-665.

Figueiredo, I. N., Leal, C., Pinto, L., Bolito, J., & Lemos, A. (2016). Exploring smartphone sensors for fall detection. *mUX: the journal of mobile user experience*, *5*(1): 1-17.

Gasparrini, S., Cippitelli, E., Spinsante, S., & Gambi, E. (2014). A depth-based fall detection system using a Kinect® sensor. *Sensors*, 14(2): 2756-2775.

Han, J., Shao, L., Xu, D., & Shotton, J. (2013). Enhanced computer vision with microsoft kinect sensor: A review. *Cybernetics, IEEE Transactions*, *43*(5): 1318-1334.

He, Y., Li, Y., & Bao, S. O. (2012). Fall Detection by built-in tri-accelerometer of smartphone. *Proceedings of Biomedical and Health Informatics (BHI), 2012 IEEE-EMBS International Conference,* 184-187.

Igual, R., Medrano, C., & Plaza, I. (2013). Challenges, issues and trends in fall detection systems. Biomedical engineering online, 12(1): 1.

Mastorakis, G., & Makris, D. (2014). Fall detection system using Kinect's infrared sensor. *Journal of Real-Time Image Processing*, *9*(4): 635-646.

Miao, F., He, Y., Liu, J., Li, Y., & Ayoola, I. (2015). Identifying typical physical activity on smartphone with varying positions and orientations. *Biomedical engineering online*, 14(1): 32.

Mubashir, M., Shao, L., & Seed, L. (2013). A survey on fall detection: Principles and approaches. *Neurocomputing*, 100: 144-152.

Sathyanarayana, S., Satzoda, R. K., Sathyanarayana, S., & Thambipillai, S. (2015). Vision-based patient monitoring: a comprehensive review of algorithms and technologies. *Journal of Ambient Intelligence and Humanized Computing*, 1-27.

Stone, E. E., & Skubic, M. (2015). Fall detection in homes of older adults using the Microsoft Kinect. *IEEE journal of biomedical and health informatics*, 19(1): 290-301.

Wang, L., Qiao, Y., & Tang, X. (2013). Motionlets: Mid-level 3d parts for human motion recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2674-2681.

Yang, L., Ren, Y., & Zhang, W. (2016). 3D depth image analysis for indoor fall detection of elderly people. *Digital Communications and Networks*, *2*(1): 24-34.