Monitoring Hospital Rooms for Safety Using Depth Images

Tanvi Banerjee[1], Marilyn Rantz[2], Mengyuan Li[1], Mihail Popescu[3], Erik Stone[1], Marjorie Skubic[1], Susan Scott[2]

[1]Electrical and Computer Engineering, University of Missouri

[2]Sinclair School of Nursing, University of Missouri

[3]Health Management and Informatics, University of Missouri

Abstract

A two-stage fall detection technique developed by our team was tested in a real hospital setting with falls acted out in a patient room. To further test the algorithm, data were collected at the University of Missouri hospital with actual patients. Features extracted from three dimensional point clouds created from Kinect depth images were used as input to the fall detection system. Kinect sensors were placed in six hospital rooms and data were collected. The data processed from the hospital setting are discussed, demonstrating the need for an automated fall detection system which has shown robustness in addressing real world challenges in a dynamic environment.

Introduction

Detection of falls is a major health problem (Sadig et al. 2004). It has been reported that 67% of fall victims who fell and remained helpless on the ground for a long duration suffered premature deaths compared to those who were detected earlier (Murphy 2000). Patient falls have devastating repercussions on the healthcare system, and a single fall can lead to reduced mobility and create a vicious cycle causing higher fall risk (Tinetti 1990).

Several fall detection systems have been created using vision sensors. Among recent work, Dubey et al. (2012) used depth information as well as color values recorded by the Microsoft Kinect sensors and computed the motion history images and the Hu moments of the RGB-D (color and depth) data. Their Support Vector Machine (SVM) classifier obtained an accuracy > 95% in a lab setting.

Other research uses visible cameras for fall detection with finite state machines (FSM). Background modeling was done to track the person in the scene; the aspect ratio was used as input to the FSM. A mixture of Gaussians was created using training images so that each Gaussian represented the aspect ratios of a particular pose (such as upright, sitting). The testing was done in laboratory settings, achieving approximately 85% success. Another system detects falls by placing the camera on the ceiling, extracting features (silhouette, lighting, and optical flow features) and training classifiers such as neural networks, SVMs and logistic regression (Belshaw et al. 2011). Best results of 92% true positives and 5% false negatives were obtained in a lab setting. Thome et al (2006) used HHMMs (hierarchical hidden markov models) for fall detection with two layers for modeling motion. The first stage distinguishes between standing and rest. Further state rules were used to trigger falls; 82% accuracy was achieved for falls and 98% for non falls. Several fall detection systems have also employed wearable sensors using accelerometers on the torso and thigh (Nyan et al, 2008), or on the trunk (Boissy et al. 2007).

Our work has been conducted in real hospital rooms using non-intrusive Kinect depth data, based on a two-stage fuzzy inference system that captures the history leading up to a fall (Anderson et al. 2009). Data have been collected continuously with hospitalized patients over several days; results are reported in subsequent sections. This paper reports the first testing of the Kinect sensors in a real hospital setting. Real-world challenges are discussed, and proposed improvements are suggested.

Fall Detection System

In our fall detection system, depth data are captured using the Microsoft Kinect, and the foreground is then extracted. A 3D point cloud is created using the depth information obtained from the sensors. Features extracted are fed to a fuzzy inference system. To address privacy concerns, we use information from the depth images only, which effectively provide a 3D silhouette.

Foreground Extraction

The Microsoft Kinect sensor uses a pattern of actively emitted infrared light to produce a depth image and allows for a 3D representation using a single Kinect. Two factors, namely its low cost and its invariance to lighting conditions make the Kinect sensors an ideal sensor for monitoring safety of patients in hospital rooms. Also, the depth images address privacy concerns of patients while still providing important activity information (Stone & Skubic 2011).

The foreground extraction is performed on the raw depth images using a dynamic background subtraction algorithm. The 3D point clouds of these extracted foreground objects, formed using the known intrinsic and extrinsic parameters of the Kinect, are then tracked over time. Features extracted from these 3D clouds are then used as input to the fall detection system. Figure 1 shows the flow of the entire algorithm.

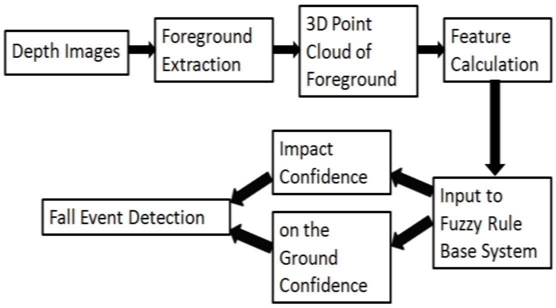

Figure 1: Block Diagram of the Fall Detection Algorithm Using a Fuzzy Rule Based System.

Fuzzy Inference System

Fuzzy inference systems interpret a set of rules expressed as fuzzy conditional statements (Mamdani 1974). Our two-stage rule-based fuzzy inference system has been adapted from the work described in (Anderson et al. 2009), which was developed using visible image data collected from two web cameras. The first layer of the classifier uses the states *Upright*, *On the Ground* and *In Between*. For this, two main features are used: the centroid height and the maximum height of the detected object. The rules for the first stage are given in Table I.

Depending on the values of the feature set at each frame, certain rules get fired and the membership values of the features in the fuzzy sets (L, M, H) fire the respective rules. The state decision at each stage is made by choosing the state with the highest membership value. If the degree of membership is low for all the states, the state will be undefined; this is part of the strength of the algorithm, since the wrong state is not selected. As can be seen from Table I, the Upright state is identified when the maximum height and average height of a person is high.

The second inference level has two states – fall and non-fall. The On the Ground state from level one triggers the fall rules for the second level. In particular, there are two rules which trigger the fall event. These are:

- *If* the On the Ground event is detected for a **long** duration, *then* a fall has occurred.
- *If* the Impact Confidence before an On the Ground event is **high**, *then* a fall has occurred.

The impact confidence is a function of the acceleration of the person download. The idea is that, if a fall occurs, there will be an abrupt drop in the downward direction. For our application, we use the minimum duration for the On the Ground state to be 5 seconds to trigger a fall event. The algorithm looks for a fall event over a window length of 6 seconds just before the On the Ground state. Ideally, if a fall has occurred, both the rules mentioned above will trigger. However, since this a dynamic environment with occlusions present, the impact confidence may not be present if an object occludes the person from the Kinect view. Hence, both the rules are kept independent. Figure 2 shows the results of the two-layer system for a given fall.

Table I. Fuzzy rules for first level of fall detection
V = very low, L = low, M = medium, and H = high.

	Centroid	Max Height		Upright	In Between	On the Ground
If	H	H	Then	H	V	V
	M	H		M	V	V
	L	H		L	L	V
	H	M		M	L	V
	M	M		L	L	V
	L	M		L	H	V
	M	L		V	H	L
	L	L		V	V	M
	H	H		L	V	V
	M	H		L	L	V
	L	H		V	L	H

Data Collection and Results

For our initial experiments, we used an empty hospital room at the University of Missouri Hospital where three of our researchers performed 18 falls (e.g., walking then falling down or falling from the bed) and 17 non-fall events (e.g., crouching down, stooping to tie shoe laces, lying on the floor) for a total duration of 44 minutes. All the falls and non-falls were correctly identified.

Figure 2 shows the results of one of the falls correctly identified by the fall detection system. The bottom graph shows the three states of the first level. As shown, whenever the membership of the person at a given time is high for the Upright state, the membership for the On the Ground state is low and vice versa. The level two results are shown in the top graph which indicates the presence of the fall with varying membership values. The fall is identified at the frames where the membership value is 1. The minimum threshold for the fall event used here is 0.5.

To collect realistic data in the hospital setting, we placed Kinect sensors in six patient rooms in the hospital, configured to log data continuously 24 hours a day. For this paper, we discuss the results of depth data collected continuously in a hospital room over a period of 17 days. During this period, no falls were reported by the hospital staff. Our algorithm reported 12 falls so these are all false positives identified by the algorithm. Figure 3 shows the Kinect (circled) attached to the wall mount TV using brackets. The sensor was placed approximately 7-9 feet from ground level.

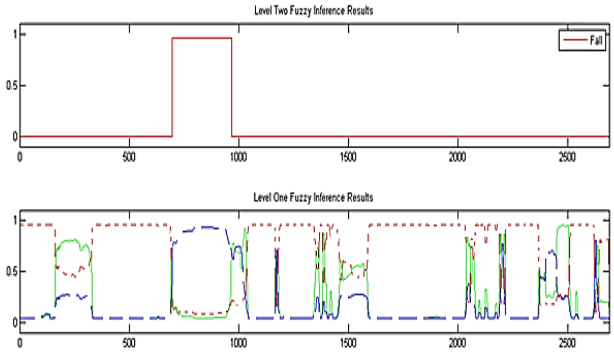

Figure 2: Results of the Fall Detection Algorithm Using a Fuzzy Rule Based System to detect a fall. The bottom graph represents the membership value of each image frame for the three states On The Ground (large dash), In Between (no dash) and Upright (small dash) whereas the top graph identifies the Fall event (membership value) which comprises the second inference level. The X axis represents the image frame number.

Figure 3: Kinect mounted below the TV in a hospital room

Figure 4 shows example depth images, the extracted foreground, and the 3D clouds created from the extracted foreground. Images a-c are of a fall by a researcher; images d-f are of an actual hospitalized patient. It can be seen from the figure that there is a lot more clutter present in the dynamic environment in the hospital room when there is an actual patient in the room due to all the health monitoring devices and equipment.

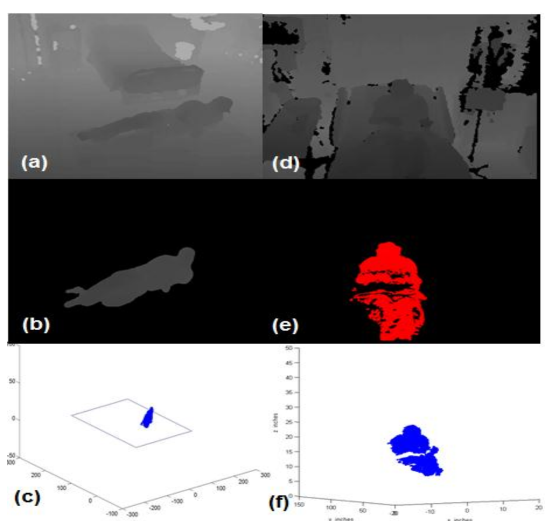

Figure 4: Depth (a,d), extracted foreground (b,e) and 3D point clouds (c,f) of a person falling (a-c) and an a hospitalized patient on the bed (d-f).

A point to note here is that hospital staff move in and out of the room freely, whereas the patients are often wheeled into the room, e.g., on a gurney and transferred to the hospital bed, which is on wheels and is sometimes moved around. These create unique challenges for automated fall detection.

Challenges and Lessons Learned

Conducting research in the hospital environment requires much pre-planning by the research team in partnership with the clinical, engineering, patient safety, risk management, infection control, regulatory compliance, and human subjects' protection staff. Installation of the equipment requires multiple layers of approval and supervision to assure all regulatory standards are met. For example, although there are ceiling tiles that could enable cable or wiring installation for placement of equipment, this approach requires additional installation costs and time due to infection control and hospital regulations.

Access to patient rooms is limited. When a patient is discharged, the room must be cleaned and quickly prepared for the next patient admission, so there is little time for installing the equipment. Methods for obtaining the data are tested before installation and confirmed as working correctly before each installation. We are using easily removable external hard drives for each room which are replaced by the clinical manager on the unit between patient admissions. This approach is working well to obtain the data for analysis. As the system is further refined, secure wireless methods will be approved by the hospital information technology staff so that real time staff notification of potential falls or increasing fall risk can be sent.

Hospital staff and patient acceptance of using images for patient safety is always a concern. When we, as clinicians and researchers, initially approached hospital clinical, patient safety, and human subject protection staff, we demonstrated the "shadow-like" data collected in an empty patient room while we discussed how the images could be helpful in learning about the movements that may precede a fall. The staff were impressed with the anonymous images and concluded these were much less privacy invading compared to traditional video used in some hospital areas for safety and security. Simple posters display a sample image and explain that hospital staff are testing new methods to keep patients safe. Thus far, no concerns have been communicated to hospital staff by patients or family.

As mentioned earlier, 12 false positives were reported in the patient room. The constant shifting of the patient bed and the movable tray and the frequent visits from the hospital staff and relatives create the clutter which is to be expected in a live setting. Also, the patient being wheeled in and out of the room using gurneys constantly displaces the background model. In spite of such a dynamic setting, we get an average false positive of less than one per day which is very promising. In addition, using a two stage classifier makes the system more transparent compared to an SVM classifier since we can see which sequences trigger the On Ground event and then lead to a fall. This gives more insight into the scenario leading up to the fall and illustrates how the classifier works. Ultimately, this will lead to better fall detection and better analysis of falls that can aid in developing fall prevention strategies.

Figure 5 shows an example of a false positive scenario. Here, the highlighted white region is identified as a fall. This foreground is caused by the movement of the tray with the highlighted region being the original location of the tray. Once the tray moved, the empty spot created by the movement triggered the On the Ground rule and wrongly identified it as a fall. All of the 12 false positives represent similar scenarios. Future work will investigate better reasoning to correct this problem.

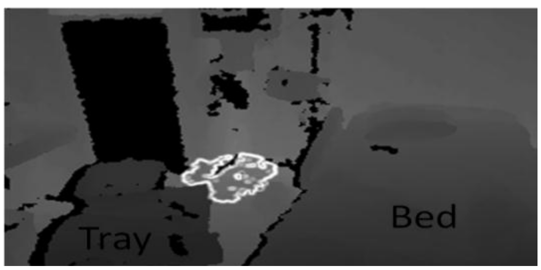

Figure 5: Example of a false positive detected by the algorithm. The highlighted region in white shows the object identified as On the Ground.

This paper reports the first real-world testing of Kinect depth images in the hospital setting for detecting falls and the conditions under which they occur. The analysis reported is from the first 17 days of data from an installation in the hospital with real patients being cared for by staff. We also report the researcher falls conducted in the empty room to pre-test our algorithms in the actual hospital environment. Although no falls have occurred in any of the six rooms where the depth images are being continuously collected, some are likely to occur within the upcoming months for analysis. We are currently working on the challenges as equipment is moved within the patient rooms, staff movement as they provide care, and quiet resting of patients in the bed. Results using our algorithm have been compared with results obtained using Hidden Markov Models for modeling body postures (Anderson 2009) and have achieved higher success rates. Future work involves adding better automated reasoning to distinguish an inanimate object from a person. This can be addressed by increasing the frequency of the dynamic background update to remove stationary objects; however, further background errors may be introduced due to the patient lying still in the bed. We are investigating features to distinguish objects from a person in such a dynamic setting. We are also adding a tracking module which is independent of the background model so that any object which is not tracked in previous frames is considered inanimate and can be discarded.

Acknowledgement

This research was funded in part by NSF grant CNS 0931607. We would especially like to thank nurse manager, Roland Ragan, RN for his assistance.

References

Anderson D. 2009. Anderson D, Luke RH, Keller JM, Skubic M, Rantz M and Aud M, "Linguistic Summarization of Video for Fall Detection Using Voxel Person and Fuzzy Logic," *Computer Vision and Image Understanding*, 113: 80-89.

Anderson D, Luke RH, Keller JM, Skubic M, Rantz M and Aud M. 2009. Modeling Human Activity from Voxel Person Using Fuzzy Logic. *IEEE Trans. on Fuzzy Systems*, 17(1): 39-49.

Belshaw M, Taati B, Snoek J, Mihailidis A. 2011. Towards a Single Sensor Passive Solution for Automated Fall Detection. *Proc*., *Intl. Conf. of the IEEE EMBS*, Boston, Aug. 30 – Sept. 3.

Boissy P, Choquette S, Hamel M, Noury N. 2007. User-Based Motion Sensing and Fuzzy Logic for Automated Fall Detection in Older Adults. *Telemedicine and e-Health*., 13(6): 683-694.

Bonner S. 1998. Assisted interactive dwelling house: Edinvar housing association smart technology demonstrator and evaluation site. *Improving the Quality of Life for the European Citizen (TIDE)*, 396–400.

Dubey R, Ni B, Moulin P. 2012. A Depth Camera Based Fall Recognition System for the Elderly. *Proc., Intl. Conf., ICIAR*.

Mamdani, E. H. 2000. Application of fuzzy algorithms for the control of a simple dynamic plant. *Proc IEEE*, 121-159.

Murphy SL. 2000 Deaths: Final data for 1998. National Vital Statistics Reports, Hyattsville (MD): *National Center for Health Statistics* 48(11).

Nait-Charif H, and Mckenna S. 2004. Activity Summarization and Fall Detection in a Supportive Home Environment. *Proc., Intl. Conf. on Pattern Recognition*.

Nyan M, Tay FE and Murugasu E. 2008. A wearable system for pre-impact fall detection. *J. Biomech*, 41(16): 3475-3481.

Sadigh S, Reimers R, Andersson R, Laflamme L. 2004. Falls and fall related injuries among the elderly: a survey of residential-care facilities in a Swedish municipality. *J. Community Health.* Apr; 29(2):129-40.

Stone E & Skubic M. 2011. Evaluation of an Inexpensive Depth Camera for In-Home Gait Assessment. *J Ambient Intelligence and Smart Environments*, 3(4):349-361.

Thome N and Miguet S. 2006. A HHMM-based approach for robust fall detection, *Proc. Intl. Conf. on Control, Automation, Robotics and Vision*, 1-8.

Tinetti ME, Richman D, Powell L. 1990. Falls efficacy as a measure of fear of falling. *J Gerontol*. 45(6):239-243.