Fall Disclosure in Homes of Older Adults Using Kinect

Sargunavathi.S[1], Harish.M[2], Jagan.D[3], Kiran.S.S[4]

Associate Professor & Electronics and communication & Sriram Engineering College (Anna University) Thiruvallur, India.

Abstract — *A module for recognizing decline in the home of older people using Microsoft Kinect and a two stage decline recognizer system is presented. The first stage of recognizing system distinguish a person's steep event in particular depth image frames, and then fragments on ground events from the steep event time series capture by tracking over time. To calculate a confidence that a fall occurs on ground event the second stage uses the ensemble of decision trees. Evaluation was conducted by using the data set consists of 454 falls performed by professional stuntman .These various data collection allows the characterization of performance of the system under real time conditions to a new level. These results including the standing, lying down, near or far fall locations. By using this method better results are achieved.*

Keywords—*Accelerator, Depth sensor, fall detection, ground event, Kinect.*

I. INTRODUCTION

Falls are a crucial concern among older adults. It is estimated that one out of three older adults (age 65 and over) falls each year. Those who fall may suffer serious injuries, like as fractures in hip and head traumas, self-reliance and guide to an increased possibility of prior death. The medical cost of falls among older people in INDIA in the year 2000 was $19 billion. This cost does not narrate for reduced quality of life and other long term problems suffered after a fall. Studies have also found the increased risk of physical and physiological complications joined with prolonged periods of lying on the floor following a fall, due to an inability to get up. Older adults living alone are at a risk of delayed assistance following a fall. A low-cost, reserved system capable of automatically, recognizing decline in the homes of older people could help drastically reduce the incidence of delayed assistance. This paper is organized in a way that it has three steps .First it interrogate with the related work and the Second states about Kinect based systems and Third is for detecting falls from the data.

II. RELATED WORKS

The existing system for detecting falls are, An Automatic fall detection using smartphone. The generated acceleration patterns from activities are analysed to detect the fall.[1].The next method involves three main categories they are wearable devices, ambient analyser and motion detector to detect the falls[2].This method involves the accelerometer sensor to detect the sudden downward tilt. So here 3 axial accelerometer is used to provide measure of acceleration in three dimension [3].This module involves Gyroscope based fall detection sensor array. These have stimulated falls to differentiate a fall and false alarm .Falls can be easily found from ADL with 100%accuracy [4].In this method involves the floor sensor based on near field imaging to detect the falls. This method uses near field imaging and pattern recognition to detect falls [5].This process introduces Human retina like optical sensor system which provides privacy to patients. This method also have pre-recorded data streams of real falls to distinguish the original falls and false fall[6].According to Robust fall detection based on spatio-temporal context switching to create 3D images. Then dense spatio temporal context is used to trace the head posture and estimate the distance to the floor[7].This module involves wrist worn detection of unexpected fall as it is easier to use and has an immediate notification call option to notify action to be taken. As the process is simple and uses three axes of acceleration sensor and make more useful to use[8].In this process uses ultra-sonic sensor with Arduino microcontroller ,here threshold signal plays a important role with room temperature changes and it can state the fall gesture with accuracy and uses two temperature sensor for anti-crash system[9].Here it is used for the patients in the hospital for the detection of an abrupt fall and here the signals are sampled to low frequency to identify an fall ,high sampled signal identify the acceleration magnitude ,the posture of patient and magnitude with three axis accelerometer signal are used to locate the fall detection[10].This process initiate a negative acceleration for a fall with use of accelerometers with body fall directions and movements[11,12].This approach scales metabolic expenditure,balance,physical activities & postural sway,gait to study a fall with accelerometer[13].Here Tunstall is used with two form, one create a wake-up call if fall is detected and another method initiate an alarm if he/she remain asleep with certain measurement[14].This project describes about the fall & medical cost of old aged adults who live alone in their house as it would affect their life by both physically & economically[15].Here the project states the complexity of prolonged time of lying on the floor after a fall, due to incapability to rise-up. As this would increase the risk of the patient life[16].In this project it overcomes the state of wearable module as it would be having more drawbacks like change of battery, lose sight and so on. So, here preferred module is non-worn sensor[17].In order to overcome wearable sensor, this method detect a fall with floor vibrating sensors &acoustic sensors with reliability[18].According to this module it uses many cameras located in a living room to obtain a three dimension image of floor objects. As it help in notifying a fall and differentiate a object & human fall [19, 20].

III.PROPOSED SYSTEM

All for fall detection a two stage process is used. The stage one of the process characterizes the three dimension object for individual frame and it finds ongoing ground events through temporary partition of vertical series of three dimension objects. The stage two of the process uses the decisions and features taken from stage one to calculate that a fall occurs

A. First Stage — Vertical State Characterization

To characterize the vertical position of an object to an single frame three features are used they are max height, centroid, total number of discrete elements. Where T_{max} is set to 170cm (average height), T_{cent} is set as 85cm (which is ½ of T_{max}),H_{pg} is found from the data. D_{pg} ranges from 350 to 391. When a person is standing or leaning the range varies from 1.6 to 2.2 and 0.9 to 1.4 if the person is on the ground after a fall the range will be zero or less

B. First Stage — On Ground Event Segmentation

The 3D object is tracked over time and estimated by vertical state. The vertical state time series temporal segmentation identifies the on ground events. The robust feature extraction is used for accurate segmentation of fall motion. The duration of fall motion vary according to the type of fall. So it can create problems in feature extraction if fixed window size is used. The smoothed signal is yielded by filtering vertical time series and average filters

Where f_{rate} is the frame rate, and T_{trig} is a threshold learned from training data as the value required to trigger$_{on}$ *ground event* extraction (if possible) for all falls, plus five percent. During evaluation, this value ranged from 0.876 to 0.903 over 13 cross validation folds. To allow real-time response to fall events, if the end of an *on ground event*, *tend*, is not identified within four seconds of the start, t_{start}, fall confidence is computed at that time, denoted t_{fend}, before *tend* is identified.

Thus, the time elapsed from t_{fall} to t_{fend} is at most eight Seconds. The point t_{init}, which denotes when *on ground event* extraction was triggered, is included for illustrative purposes.

From one stunt actor standing fall, one stunt actor sitting fall, and one naturally occurring resident fall. Finally, to address the issue of objects blending into the floor in the depth imagery from the Kinect (and thus no longer being tracked), if the time elapsed from *tstart* to the last time an object was tracked is less than four seconds, the last observation is artificially repeated to meet this condition. This allows falls to be detected at larger distances from the Kinect, where the faller may not be distinguishable from the floor in the depth imagery following the fall.

C. Second Stage – On Ground Event Features

Five features are extracted for each *on ground event* for the purpose of computing a confidence that a fall preceded it. A description of these features follows.

Minimum Vertical Velocity (MVV):

Minimum vertical velocity is computed as the minimum of Vs′ (the derivative of the vertical state time series) from *tfall* to *tstart*. This feature characterizes the downward vertical motion of falling to the ground.

Maximum Vertical Acceleration (MVA:)

Maximum vertical acceleration is computed as the Maximum of Vs″ (the second derivative of the vertical state time series) from the frame of minimum velocity to *tstart*. This feature characterizes the action of abruptly hitting the floor, stopping downward vertical motion.

Mean V_{avg}:

Mean $_{Vavg}$ is computed as the mean of V_{avg} from t_{start} to t_{fend}A higher value is indicative of bad foreground segmentation, or non-fall activities such as bending over or kneeling.

Adjusted Change in $Z_{pg}(\Delta_{pg})$:

When objects in the environment move, whatever the object was previously occluding may be identified as foreground depending on a variety of factors that influence the background model. Δ_{pg} measures the change in Z_{pg} from t_{fall} to t_{start}, accounting for the possible impact of occlusion. The presence of occlusion is assessed for each pixel, at each frame, using the change in measured depth. If the rate of change between the current and previous frame exceeds 420 cm/sec (selected as a value significantly faster than a person would ever be moving in their home), it is assumed that whatever the pixel is currently viewing was occluded by a different, closer object in the prior frame. Given this assessment of occlusion, the number of foreground pixels belonging to a 3D object that have a height (after projection to world 3D) below 38 cm and were previously occluded, $p_{occlude}$, can be counted at each frame, along with the total number of pixels belonging to the object that have a height below 38 cm, p_{ground}. The percentage of p_{ground} pixels appearing due to occlusion from *tfall* to t_{start} is approximated.

A near zero value of Δ_{pg} implies there was no significant change in Z_{pg} that was not a result of occlusion. A higher value is indicative of an object moving from a more off the ground position to a more on the ground position.

Minimum Frame-to-Frame Vertical Velocity (MFFVV):

Minimum frame-to-frame vertical velocity is an estimate of the minimum vertical velocity from *tfall* to *tstart* computed by mapping foreground pixels in a depth image frame, to foreground pixels in the previous depth image frame, independent of the tracking algorithm. At each frame, a pointwise matching is found (after projection to 3D) between all the foreground pixels in the current frame and all the foreground pixels in the prior frame such that the overall change in 3D position is minimized. A simple heuristic and random assignment is used to initialize the matching. The matching is then iteratively refined using a neighbourhood guided search, where neighbourhood structure is obtained from image space. A many-to-one mapping of current to previous foreground pixels is allowed as needed. After refinement, the vertical movement of a foreground object between the previous and current frame is estimated as the median of the vertical movement of all the pixels belonging to the object. MFFVV is computed as the minimum of these frame-to-frame vertical velocity estimates for an object from *tfall* to *tstart*. In the absence of tracking errors, MFFVV will likely not be as accurate as MVV derived from the vertical state time series. However, MVV will reflect any vertical motion inferred by the tracking algorithm, whether that motion is correct, or the result of a tracking error. MFFVV, meanwhile, is computed from vertical velocity estimates made at each frame independent of the tracking algorithm. Thus, MFFVV will generally be more robust when the scene contains multiple foreground objects (or foreground segmentation errors) that are in close proximity.

D. Second Stage – Ensemble for Fall Confidence

A fall is computed for each ground event using the five extracted features and an ensemble of decision trees. Each tree is a binary tree in which the decisions are based on a single, randomly selected, predictor (dimension). The tree is built recursively, in a top-down fashion, with the optimization criterion for selecting split cut points being the mean-squared error (MSE) of predictions vs. the training data. Leaf nodes continuously divide until a node remains less than 10 observations, as split can create child node with fewer than 5 observations, or the MSE for the node's predictions drops below a threshold ε. After the tree is built, the predicted value of each leaf node is the average target value of the training observations assigned to the node. Each decision tree is trained using a sample of the training data. A large imbalance, roughly 1:400, of positive (fall) to negative (non-fall) samples exists in the collected data (described in Section V). Thus, each sample of the training data contains all of the minority class examples (falls) and a randomly selected, with replacement, under-sampling of the majority class examples (non-falls). Additive Gaussian noise (zero mean with standard deviation equal to 5% of that present in the training data positive samples), or jitter, is added to each training sample to further improve generalization. The output of the trees is combined by averaging.

E. Ground Truth Matching

Matching of *on ground events* (with temporal segmentation *tfall*, *tstart*, *tfend*, *tend*) to ground truth, a set of times, $G = \{g1, \ldots, gn\}$, for each known fall, is assessed by the condition:

$$T\,fall - \varepsilon \leq g\,i \leq t\,end + \varepsilon$$

where ε is two seconds. If true, for exactly one ground truth entry, then the *on ground event* is said to match known fall *i*. The time of each known fall was identified, based on review of the depth imagery, as the time at which the faller was on the ground and vertical movement due to the fall had ceased.

References

[1] Tran Tri Dang, Hai Truong, Tran Khanh Dang, "Automatic Fall Detection using Smartphone Acceleration Sensor" (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 7, No. 12,pp.123-129,2016.

[2] Goh Yongli, Ooi Shih Yin and Pang Ying Han, "State of the Art: A Study on Fall Detection" World Academy of Science, Engineering and Technology 62,pp.294-298, 2012.

[3] Sharwari Kulkarni*, MainakBasu, "A Review on Wearable Tri-Axial Accelerometer Based Fall Detectors" Journal of Biomedical Engineering and Technology, Vol. 1, No. 3, pp.36-39,2013.

[4] A.K. Bourke, G.M. Lyons, "A threshold-based fall-detection algorithm using a bi-axial gyroscope sensor" University of Limerick, Limerick, Ireland,pp84-90,Received 5 September 2006; received in revised form 28 November 2006; accepted 3 December 2006.

[5] Henry Rimminen, JuhaLindstr¨om, MattiLinnavuo, and RaimoSepponen,"Detection of Falls Among the Elderly by a Floor Sensor Using the Electric Near Field" IEEE Transaction On Information Technology In Biomedicine,VOL.14, NO. 6,pp.1475-1476,2010.

[6] ÁgostonSrp, Dr. FerencVajda," Fall Detection System for Older People" Budapest University of Technology and Economics 1 'Eurostat 2008, EUROPOP2008, convergence scenario'.

[7] Lei Yang 1, Yanyun Ren 1, Huosheng Hu 2 and Bo Tian 3,* , " New Fast Fall Detection Method Based on Spatio-Temporal Context Tracking of Head by Using Depth Images" Sensors15,pp. 23004-23019,2015.

[8] Thomas Degen, Heinz Jaeckel, Michael Rufer, Stefan Wyss, "SPEEDY: A Fall Detector in a Wrist Watch" Swiss Federal Institute of Technology Electronics Laboratory 8092 Z¨urich, Switzerland {degen,jaeckel}@ife.ee.ethz.ch.

[9] ChokemongkolNadee, KosinChamnongthai,"Ultrasonic Array Sensors for Monitoring of Human Fall Detection" King Mongkut's University of Technology Thonburi Bangkok, Thailand Chokemongkol.nad@mail.kmutt.ac.th,978-1-4799-7961-5/15.

[10] Patrick Scholten,.Lettele (NL); Harry BernardllsAntonlusKerver, Duwen(NL),"FALL DETECTION ALGORITHM UTILIZING 5,865,760 A 2/ 1999 Lidman et al.A THREE_AXIS ACCELEROMETER"US pattern Scholten et al,Us8,814,81 1 B2,Aug.26,2014.

[11] M. Mathie, J. Basilakis, and B.G. Celler, "A system for monitoring posture and physical activity using accelerometers," 23rd Annual Int. Conf. IEEE Engineering in Medicine and Biology Society, 2001.

[12] R. Salleh, D. MacKenzie, M. Mathie, and B.G. Celler, "Low power tri-axial ambulatory falls monitor," Proc. 10th Int. Conf. on Biomedical Engineering, 2000.

[13] M.J. Mathie, A.C. Coster, N.H. Lovell, and B.G. Celler, "Accelerometry: providing an integrated, practical method for long-term, ambulatory monitoring of human movement," Physiol. Meas.25(2):R1- 20,2004.

[14] K. Doughty, R. Lewis, and A. McIntosh, "The design of a practical and reliable fall detector for community and institutional telecare," J. Telemed. Telecare 6 S150-4,2000.

[15] J.A. Stevens, P.S. Corso, E.A. Finkelstein, and T.R. Miller, "The costs of fatal and non-fatal falls among older adults," Injury Prevention, 12(5):290-295, 2006.

[16] M.ETinetti, W.L. Liu, and E.B. Claus, "Predictors and prognosis of inability to get up after falls among elderly persons," JAMA: the journal of the American Medical Association, 269(1): 65-70, 1993.

[17] G. Demiris et al., "Older Adults' Attitudes Towards and Perceptions of Smart Home Technologies: A Pilot Study," Informatics for Health and Social Care, 29(2):87-94, 2004.

[18] Y. Zigel, D. Litvak, and I. Gannot, "A method for automatic fall detection of elderly people using floor vibrations and sound—Proof of concept on human mimicking doll falls," IEEE Trans. Biomed. Eng., 56(12):2858–2867, 2009.

[19] D. Anderson, R.H. Luke, J. Keller, M. Skubic, M. Rantz, and M. Aud, "Linguistic summarization of activities from video for fall detection using voxel person and fuzzy logic," Computer Vision and Image Understanding, 113(1):80–89, 2009.

[20] E. Auvinet et al., "Fall detection with multiple cameras: An occlusion resistant method based on 3-d silhouette vertical distribution," IEEE Trans. on Info. Tech. in Biomedicine, 15(2):290-300, 2011.