Young Children's Fall Prevention based on Computer Vision Recognition

Hana Na, Sheng Feng Qin, David Wright
School of Engineering and Design
Brunel University
Uxbridge, Middlesex, UB8 3PH
UNITED KINGDOM

Abstract: - In this paper a computer vision system is proposed to detect risk factors of young children's falls in the home environment and to produce actions to remove the factors. The system recognition tasks, clutter detection and children tracking, are defined in accordance with general suggestions which request a caregiver's continuous supervision to prevent young children's falls from the UK Child Accident Prevention Trust (CAPT). The current system uses only one commercial camera without any sensor or marker on the subject for practical purposes. This paper focuses on the system design and clutter detection. The algorithms for moving object and clutter detection have been developed, implemented and tested.

Key-Words: - Fall prevention, Risk factors, Background subtraction, Motion detection, Clutter detection

1 Introduction

According to the UK Child Accident Prevention Trust (CAPT), every year over two million children are taken to hospital due to accidental injuries, and around half of these accidents happen at home [1]. Falls account for over 40% of all home accidental injuries of children, and young children aged under five are most vulnerable to injuries in the home environment where they spend most of their time [2].

As young children are not able to assess risks for themselves, the best way to prevent their falls would be parents' continuous supervision and instruction. But parents or caregivers cannot keep an eye on their children all the time. Therefore, a computer-vision-based system is proposed in this paper to keep watch on children at home and alert their nearby parents or caregivers when it finds fall-potential situations.

Many applications have been developed to detect falls of the elderly which could be fatal [3][4][5][6][7][8][9][10]. They use acceleration sensors to be worn by users and cameras individually or together for the fall detection. Although some of them collect the fall data from the sensors to evaluate the user's personal fall risks for later prevention, there is no prevention against falls during the data collection and against irregular falls even afterwards. Some wearable devices provide a prompt protection such as an airbag and an overhead tether when sensing a fall, but the user should wear it all the time.

The system proposed here uses only one fixed web-camera to detect risk factors of young children's falls in the home environment and give a caregiver an alert to get rid of the factors or directly guide the subject children before a fall happens.

For the fall risk factors of elderly people, there are generally intrinsic factors such as chronic diseases, cognitive impairment and sensory deficits, and extrinsic factors including environmental hazards (e.g., slippery surfaces) or hazardous activities (e.g., inattentive walking) [11]. As intrinsic factors are about health problems, most of young children's falls would be based on the extrinsic factors which are related to their environment. Therefore, the young children's fall prevention should focus on the young children's environmental and behavioural factors related to falls. This is the focus for our research.

CAPT suggests several tips to prevent falls of babies from birth to toddling [12]. The tips requesting parents' constant supervision rather than environmental instant modification were extracted from them as below;

- keep floors clear of toys and other clutter
- make sure there are no sharp edges that could cause injuries when they fall
- make sure there is no furniture around they can climb

Based on the above tips and the general knowledge that children easily fall in running, our system tasks are identified as the following;

- check any clutter's appearance on the floor

- check if the children run too fast
- check if the children move close to any structure in their environment
- check if the children climb any furniture

From the perspective of computer vision, the above tasks could be divided into clutter detection and tracking of moving children, and this paper only presents the clutter detection.

2 Related Work

The proposed system should watch both of small objects and human beings to detect any clutter's appearance on the floor and children's fall-prone behaviours. Therefore, a human being and an object must be distinguished accurately using not only motion but also other reasonable cues. This section gives a brief overview of existing works to detect clutter and human beings based on different cues and track them individually.

Lipton et al [13] detect moving targets by using the pixel wise difference between consecutive image frames and classify them into human, vehicle or background clutter, based on the target size and the shape dispersedness as humans are smaller than vehicles and have more complex shapes. This method is somewhat simple which is good for real-time motion analysis, however it seems only good enough to distinguish humans from big vehicles and the tiny motion of trees.

VIGOUR of Sherrah and Gong [14] finds skin colour clusters and tracks three boxes respectively bounding a head and two hands for one person. The head box tracker is initialised using Support Vector Machine face detection and the hand box trackers are initialised heuristically with respect to the head position for tracking of multiple people. VIGOUR also uses a simple method using a colour cue, but the subjects should be initially facing the camera and the faces should not be occluded.

The single view tracking of Cai and Aggarwal [15] is composed of background subtraction, human segmentation and the human feature correspondence between adjacent frames. After the background subtraction, human and non-human moving regions are distinguished using moment invariants based on Principal Component Analysis (PCA). And location, intensity and geometric information of the human are extracted for the tracking. The use of the three features to track human achieves much better tracking than the use of any individual feature, but the occlusion is a major obstacle.

After background subtraction, Schleicher et al [16] use a Particle Filter (PF) algorithm to identify and individually track any moving objects, and apply PCA to the each object to classify into person or non-person using geometrical constraints of several body parts. This system is relatively robust at occlusion but long-term occlusion and the lateral view of a person cause some failures.

Micilotta [17] also uses PF to track each human body after fitting a torso primitive to foreground regions and segmenting skin tone regions for the face and hands. Meanwhile, he presents a more robust method of tracking a human. Body part detectors trained by AdaBoost, detect several body parts by using skin colour cues to reduce false detections, and RANSAC assembles the parts into body configurations.

The more cues are used to detect and track human beings, the more accurate the results are. However, the use of many cues or complex methods would require expensive computation and take too long for real-time applications.

3 System Overview

The objectives of clutter detection are to find standstill objects on the floor which might trip children and to find objects which are thrown towards and hit children. The current system detects any appearance of clutter and just discriminates between standstill objects on the floor and moving objects in the scene. The information about the moving objects will be used to check if they are moving towards children after the future work, tracking of children is done.

The clutter detection consists of background subtraction, updating background image and motion detection to detect and classify clutter, and respective action against moving and standstill objects. The system workflow is presented below.

Firstly, background subtraction detects anything on the floor by comparing current images and an initial background image containing a clear floor. To focus on the floor area, non-floor area is masked in the background image.

The background image needs to keep being updated due to extraneous changes like the illumination variance. There might be small motions in the background, but they are not general in the indoor environment like swaying branches outdoors

and the system only deals with lighting changes. The slight changes of the sun light are discarded using thresholding and the dramatic changes in switching on/off a lamp are judged as an overall change.
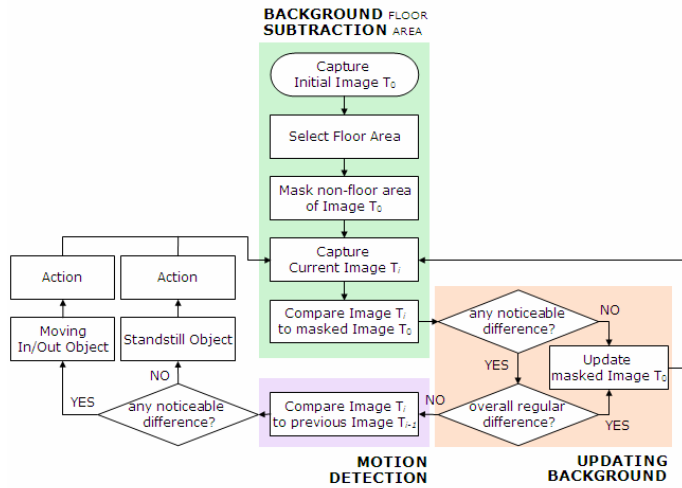

Fig. 1 System Work Flow

In the detection of any standstill objects on the floor, background subtraction is not enough because it detects both standstill and moving regions if they do not belong to the background image. So to classify standstill and moving regions, comparison of consecutive frames to detect motion is used.

Lastly, the system produces different actions in detection of still objects on the floor and moving objects in the scene.

4 Implementation
A single Logitech Quickcam Pro4000, which has the highest performance among commercial CCD webcams, is used to capture real-time images. The image size is 320x240 pixels and the developed software has dialog-based interfaces to set up and control the system.

Sub-tasks of clutter detection using image processing are floor selection, background subtraction including background update and motion detection. The details of the tasks and actions after recognition are described in this section.

4.1 Floor Selection
To focus on the floor area for detecting clutter, a mask image to indicate the floor is necessary. As a fixed camera is used here, the floor detection is required

only one time, when the camera is set up, and the software lets the user select the floor region in the initial background image.
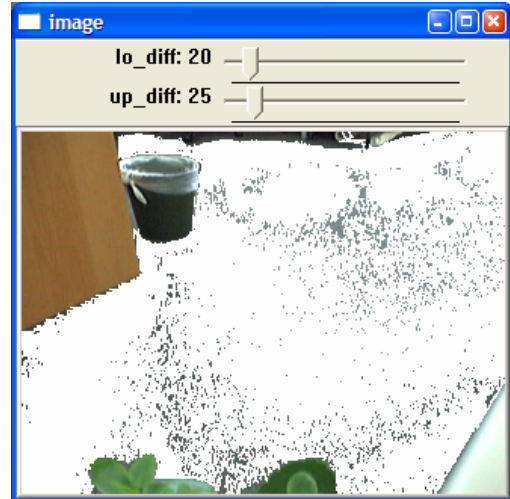

Fig. 2 FloodFill

After capturing a background image, a window of the image (Fig.2) pops up. In this window, FloodFill fills neighbour pixels whose values are close to the pixel clicked by the user. The pixel will belong to the repainted domain if its value v meets the following condition;

$$v_0 - d_{lw} \le v \le v_0 + d_{up} \qquad (1)$$

where v_0 is the value of one of the pixels in the repainted domain beginning with the clicked pixels [18]. d_{lw}, the maximal lower difference and d_{up}, the maximal upper difference between the pixels, can be defined by the user as in Fig.2, and the user can select the floor area with several clicks.

Fig. 3 Floor Mask Image

As the selected area gets lots of tiny chinks, when the user submits the floor-selected image, the system finds the contours of the area and fills them in the mask image like Fig.3.

4.2 Background Subtraction

Background subtraction is a way to find the difference between current images and the background image. In this system, an initial background image with nothing on the floor is captured to consider a noticeable difference from current images as clutter.

Firstly, a simple background model is built up by accumulating several dozens of frames (N) and calculating the mean of each summed pixel value ($bgSum_{(x,y)}$) to get mean brightness [18].

$$bgMean_{(x,y)} = bgSum_{(x,y)}/N \qquad (2)$$

Then, the absolute differences ($diff_{(x,y)}$) between the background model and the current image ($Cur_{(x,y)}$) are calculated after non-floor region is masked in the both images, using the floor mask image built previously.

$$diff_{(x,y)} = abs(bgMean_{(x,y)} - Cur_{(x,y)}) \qquad (3)$$

To get rid of noise, differences smaller than a threshold value are returned to 0, and a binary image is created by returning the others to 255.

$$diff_{(x,y)} = 255, \ if \ diff_{(x,y)} > threshold \qquad (4)$$
$$0, \ otherwise$$

Whenever this binary image becomes null, the background model is updated to cope with slight lighting changes which are ignored by thresholding. For the dramatic lighting changes by turning on/off a lamp, the background model is also updated when the difference is averagely general.

4.3 Motion Detection

As background subtraction finds both standstill and moving objects, the system also detects the difference between successive frames to only find motion.
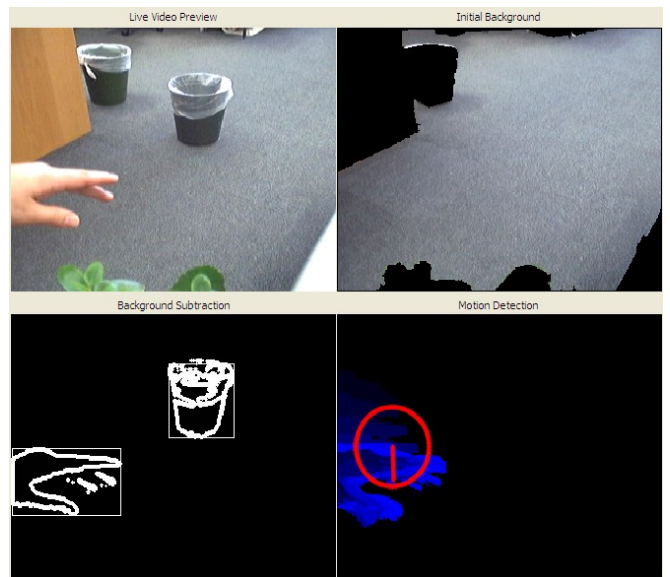

Fig. 4 Standstill vs Moving objects

In Fig.5, a slipper and a hand both appear on the background-subtracted image (left-bottom), but the difference between successive frames shows only the moving hand (right-bottom).

For the cases that there are several objects on the floor at the same time, a bounding box of each noticeable region subtracted from the background image is set as Region of Interest (ROI), and the each ROI is checked if there is any motion inside. If there is any ROI without any motion, the system considers it as still clutter on the floor. And any noticeable motion is considered as a moving object.

4.3 Action

When the system detects any still clutter on the floor, it produces a voice alert, "There is a clutter on the floor. Please move it from the floor". And the voice tells "There is a moving object" for any moving clutter.

5 Result

The floor selection works well no matter how many separate regions correspond to the floor in the background image. If there is more than one individual floor region, the contour of the each region is detected and filled respectively. As the floor is detected only once in the beginning, if any structure in the room moves in the middle of the clutter detection, the floor mask image should be updated manually. The tiny motion of a structure in the environment, however,

would not be considered as a clutter because small differences would not be set as ROIs.

As background subtraction and motion detection both compare pixel values, if the colour and texture of the clutter are very similar to the floor, the system is unlikely to consider it as clutter. So it is assumed that there is no clutter with the floor's colour and texture. Also, as the background image is updated only when the dramatic lighting changes are overall in the scene, it is assumed that there is no small lamp, only a ceiling fixture lighting the whole room.

The clutter detection works well unless a ROI is occluded by a moving region. As the bounding boxes are used as ROIs to check motion inside, if a standstill object's bounding box overlaps a moving object too much, the standstill object would be defined as a moving object. Usually the occlusion vanishes soon because one keeps the position and the other is moving. As soon as this happens, the still object is detected and the warning sound is produced.

6 Conclusion

This paper focuses on detecting risk factors of young children's falls in the home environment, to prevent falls. This is different from the previous papers that focus on detecting the actual falls and the target subject is elderly people. The risk factors are determined as environmental and behavioural ones, which are dynamic and require a caregiver's constant supervision. They include any appearance of clutter on the floor, whether the children are running fast, and whether children approach, or climbs, any structure from which they may have suffer an injury if they fall. Based on these, the tasks of the proposed system are identified as clutter detection and tracking of moving children. This paper only presents the clutter detection.

A single commercial camera is used to detect clutter for practical use without any sensor or marker to be attached on body. The clutter detection works well unless too much occlusion happens between a still object's bounding box and a moving region. The action of the system when it finds any clutter is to produce a warning voice alert at present, but it could be modified to send a message to a caregiver's or parents' portable device.

The second task of tracking moving children is being planned to check if they approach or climb furniture by measuring their position, trajectory and velocity. As the task does not require understanding of the children's posture, the method should be robust enough to classify moving regions into human or non-human and obtain accurate tracking data. Also it should be efficient and fast enough to run in real time. As the subject is young children, the classifying and tracking of them could use different features from those proposed in former papers targeting adults.

References:
[1] Child Accident Prevention Trust, Factsheet: Home Accidents, *http://www.capt.org.uk/pdfs/factsheet home accidents.pdf*, United Kingdom, 2004
[2] E. Towner, T. Dowswell, C. Mackereth and S. Jarvis, What Works in Preventing Unintentional Injuries in Children and Young Adolescents? An Updated Systematic Review*, London Health Development Agency*, 2001
[3] D. Colvin, C. Lord, G. Bishop, T. Engel and A. Patra, A Fall Intervention/Mobility Aid System for Elderly and Rehabilitative Populations, *Annual Internaltional Conference of the IEEE Engineering in Medicine and Biology Society*, Vol.13, No.4, 1991, pp. 1936-1937
[4] G. Williams, K. Doughty, K. Cameron and D. Bradley, A Smart Fall & Activity Monitor for Telecare Applications, *20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Vol.20, No.3, 1998, pp. 1151-1154
[5] T. Tamura, T. Yoshimura, F. Horiuchi, Y. Higashi and T. Fujimoto, An Ambulatory Fall Monitor for the Elderly, *22nd Annual EMBS International Conference*, 2000, pp. 2608-2610
[6] K. Fukaya, Fall Detection sensor for Fall Protection airbag, *Annual Conference of The Society of Instrument and Control Engineers*, 2002, pp. 419-420
[7] N. Noury, A Smart Sensor for the Remote Follow Up of Activity and Fall Detection of the Elderly, *2nd Annual International IEEE-EMBS Special Topic Conference on Microtechnologies in Medicine & Biology*, 2002, pp. 314-317
[8] B. Najafi, K. Aminian, F. Loew, Y. Blanc and P. Robert, Measurement of Stand-Sit and Sit-Stand Transitions using a Miniature Gyroscope and its Application in Fall Risk Evaluation in the Elderly, IEEE Transactions on Biomedical Engineering, Vol.49, No.8, 2002, pp. 843-851
[9] A. Sixsmith and N. Johnson, A Smart Sensor to Detect the Falls of the Elderly, *IEEE Pervasive Computing*, Vol.3, No.2, 2004, pp. 42-47

[10] H. Nait-Charif and S. McKenna, Activity Summarisation and Fall Detection in a Supportive Home Environment, *17th IEEE International Conference on Pattern Recognition*, 2004

[11] K. Perell, A. Nelson, R. Goldman, S. Luther, N. Prieto-Lewis and L. Rubenstein, Fall risk assessment measures: and analytic review, *Journal of Gerontology: Medical Sciences*, vol.56, no.12, 2001

[12] Child Accident Prevention Trust, Factsheet: Falls in the Home, *http://www.capt.org.uk/pdfs/ factsheet falls.pdf*, United Kingdom, 2004

[13] A. Lipton, H. Fujiyoshi and R. Patil, Moving Target Classification and Tracking from Real-Time Video. *IEEE Workshop on Application of Computer Vision*, 1998, pp. 8-14

[14] J. Sherrah and S. Gong, VIGOUR: A System for Tracking and Recognition of Multiple People and their Activities, *15th International Conference on Pattern Recognition*, Vol.1, 2000, pp. 179-183

[15] Q. Cai and J. Aggarwal, Tracking Human Motion in Structured Environments using a Distributed-Camera System, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.12, No.12, 1999, pp.1241-1247

[16] D. Scholeicher and L. M. Bergasa, People Tracking and Recognition using the Multi-Object Particle Filter Algorithm and Hierarchical PCA Method. *EUROCON 2005 - The International Conference on "Computer as a tool"*, 2005

[17] A. S. Micilotta, Detection and Tracking of Humans for Visual Interaction, *submitted for the Degree of Doctor of Philosophy from University of Surrey*, 2005

[18] Intel Corporation, Open Source Computer Vision Library: Reference Manual, *http://prdownloads.sourceforge.net/opencvlibrary /OpenCVReferenceManual.pdf?download,* 2000