Preliminary investigations into human fall verification in static images using the NAO humanoid robot

E.Catley, K. Sirlantzis, G. Howells, S. Kelly
Image and Information Engineering Research Group
School of Engineering and Digital Arts, University of Kent
Canterbury, UK.
Email: EC364@kent.ac.uk

Abstract. Advances in medicine have led to an ever aging population who generally wish to retain a significant amount of independence. To this end there has been significant research into technologies that allow for this level of independence while maintaining an appropriate level of safety. One of the most significant risks to the elderly is the danger of falling and in light of this fall detection and alarm systems have been the focus of much of this research. This technology has often been resisted by those it is trying to help. Failing to strike the balance between several factors including reliability, complexity and invasion of privacy has been prohibitive in the adoption of this technology. Whereas some systems rely on cameras being mounted so as to allow complete coverage of a user's home, others rely on being worn 24 hours a day; this paper explores a system using the mobile humanoid NAO robot to perform fall detection.

I. Introduction

A report recently stated that in the over 65s falls are the leading cause of both fatal and non-fatal injuries [1] and that one in three of this population fall every year. This population continues to increase with the number in the UK alone expected to double to over 20 million by 2050.

As part of the COALAS (Interreg IVA) project, work has been undertaken into providing minimally intrusive in home care for the elderly and disabled, with one of the aspects being fall detection. Other literature in the field has explored a variety of solutions to this problem, such as wearable technology implementing accelerometers and gyroscopes, multiple ceiling mounted camera systems for house-wide coverage, or single camera systems for localised coverage, these implementations generally suffer from certain drawbacks.

Wearable technology exists generally falls into two categories, sometimes with overall. These are fall alarms and/or detectors. In the case of the former this is only capable of sending out a manually activated distress signal which after a fall, depending on the seriousness, the user may be incapable of activating. In both cases one of the most limiting factors is the intrusiveness of the system. Studies have found that the device is not worn at all times either deliberately, due to the discomfort brought, or unintentionally if forgotten after having been temporarily removed [2].

More recently there has been research into systems using either one or more cameras and image processing algorithms. These systems have benefitted from being relatively non-invasive but have problems related to limited field of view, reliability and acceptance among the target demographic.

This paper details the implementation of the second half of a system introduced in the paper [3]. This second half consists of the fall verification stage which utilises the robots mobile nature to intelligently move towards the potentially fallen patient and identify whether a fall truly has occurred or not.

II. Related work

The term fall verification is used throughout this paper. In this context the distinction from fall detection is that the fall verification occurs on static scenes containing fallen people whereas fall detection takes place on sequences where the person or region of interest is moving.

The majority of work in this field relies on posture estimation to determine whether the person in question is in a typical or atypical position. Some of the difficulties that face these approaches have been incomplete images of the fallen, either through environmental occlusion or bodily occlusion, background models that don't remain static due to moved furniture etc., and processing time required for the image processing.

To date the majority of work pertaining to posture estimation of static images has focussed around upright humans. Methods proposed in this regard include [3] [4] [5] While these approaches have had some success with humans in an upright position, the task of applying these methods to horizontal humans with images suffering from the previously mentioned issues is non-trivial.

The methodology used in the single wide angled lens camera approach of [6] results in a false positive rate of 0.31 and false negative of 0.22. This method uses the difference in angle of the body axes between a standing person which, due to the nature and position of the camera, will be pointing towards the centre of the image and that of a fallen person. The limitations of this system are that the difference in angle between the body axes of standing and lying must be $\geq 28°$ and it cannot distinguish between people lying on a sofa for example, and the floor.

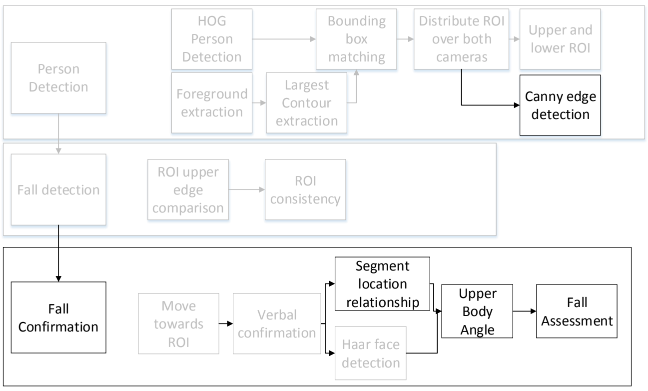
The most pertinent paper [7] which builds upon Ferrari et Al's pipeline by adding an initial estimator for limb location and sizes, thus significantly reducing the search space (the inhibiting factor of Ferrari et Al's implementation) and thus speeding computation. An average precision of 0.74 was achieved. [8]

To conclude, the biggest drawback of all these systems lies in the availability of viable images, adaptability for pose estimation in horizontal bodies and excessive computation time. Most systems make use of fixed cameras with neither prior knowledge of the approximate body location nor the ability to reposition to obtain more viable images.

III. Methodology

The overall fall detection system being implemented here falls into three main stages; Person detection, preliminary fall detection and fall confirmation. In this paper we present the fall confirmation stage, the previous two stages having been presented in previous work. While the fall detection phase left the fall detection criteria purposefully wide, achieving a high sensitivity with a low processing time at the cost of a low specificity, the fall confirmation stage is significantly slower, compromising speed for specificity. This is permissible because the fall verification routine will only run on the occasions that a fall is suspected leaving the NAO free most of the time to perform other tasks. Figure 1 illustrates the structure of this system showing the sections covered in this paper (dark) and those either left for future work or already covered in previous research (light).

Figure 1



1) *Move Towards Region of Interest*

The NAO benefits from an already implemented walking algorithm. The robot knows in which direction the potentially fallen person is and can plot a course towards them. By doing so the NAO is capable of obtaining images from several different angles until an accurate assessment can be made. Due to the current unreliability of these walking algorithms however, this stage has not been implemented and the NAO is instead moved manually to the required locations.

2) *Attempt verbal confirmation*

The microphones on the NAO and its in built algorithms for speech recognition are available. When restricted to simple one word answers such as "yes" and "no" this yields a good sensitivity and specificity making it a fast and reliable method for determining if a fall *hasn't* taken place. The reason this can only be used for reliable true negative confirmation is that in the event of a fall the user may be incapacitated and unable to answer.

This stage has been left out of scope of this paper since it is very difficult to form a test where this can be included. What proportion, for example, of falls would a subject be able to answer from? How clearly would one answer even if one were able to? Would it be the more difficult to visually classify falls that a user wouldn't be able to answer after? For the sake of maintaining scientifically sound test data this has been omitted.

3) *Transected image Canny based body identification*

By performing canny edge detection on the ROI generated during the person detection phase followed by taking a vertical slice directly through the centre of the image, it is possible to form a description of the person of interest, using the ratio of the distances between edges. By using the ratio rather than absolute values this allows us to describe this person in a scale invariant way.

Canny detection is performed once again on the image in which we are trying to verify the fall but this time horizontal slices are taken and the ratio of edge distances once again calculated. By comparing the ratios of these two images it is possible to determine whether the person happens to be in a horizontal or near horizontal position. With just three points being matched it is possible to determine the direction of the fallen body.

Fig2 shows the transition from the input images to the Canny filtered with the vertical and horizontal strip with the detected edges circled.

Figure 2

4) *Image Rotation*

To extract the person from this foreground image the largest contour is found, in almost all cases the largest contour is the subject that we wish to track. Save for other people or pets, there is unlikely to be any other foreground objects and thus few other contours.

5) HAAR face, profile face and upper body detection

For each rotation step the image will be passed through three HAAR cascades which attempt to identify the face (both frontal and profile) as well as the upper body. These features were chosen due to their ability to infer the body's angle relative to the ground from a variety of viewing angles. By calculating the approximate head position from step 3 this stage which computationally taxing can be reserved for use in locations where results are likely to be found.

6) Fall assessment

Based on the predicted angle of the body relative to the ground and verbal confirmation, an assessment is made as to whether a fall has occurred or not.

IV. RESULTS

While no true tests have yet been performed initial results are encouraging. A total of three preliminary data sets were captured with the NAO, each comprising of one reference image of the person, captured during person detection, followed by images of a prone figure captured from both front and rear. On all three of these sets a correct determination of the figure being horizontal was made.

V. DISCUSSIONS AND CONCLUSIONS

This paper has introduced the preliminary work into using edge detection with a reference and sample image to determine whether a fall has occurred in a system involving the NAO humanoid robot. While no thorough tests have been conducting making comprehensive analysis at this stage impossible, the initial results have proven the concept.
Future work will comprise of a series test of the entire system which will require the generation of a dataset, capable of evaluating each individual part; person detection, fall detection and fall verification, as well as the performance of a system as a whole.

VI. ACKNOWLEDGEMENTS

This research is financed by the INTERREG IVA France (Channel) – England European cross-border co-operation programme, which is co-financed by the ERDF.

VII. REFERENCES

[1] WHO, “WHo Global report on falls Prevention in older Age,” WHO, 2007.

[2] Senior Watch, “A market study about the specific IST needs of older and disabled people,” 2001.

[3] E. Catley, K. Sirlantzis , S. Kelly and G. Howells, “Non-overlapping dual camera fall detection using the NAO humanoid robot,” in *Fifth International Conference on Emerging Security Technologies*, Madrid, 2014.

[4] M. Marin-Jimenez, F. Vittorio and A. Zisserman, “Progressive search space reduction for human pose estimation,” in *Computer Vision and Pattern Recognition*, 2008.

[5] A. Mykhaylo, S. Roth and B. Schiele, “Pictorial structures revisited: People detection and articulated pose estimation,” in *Computer Vision and Pattern Recognition*, 2009.

[6] M. W. Lee and I. Cohen, “Proposal maps driven MCMC for estimating human body pose in static images,” in *Computer Vision and Pattern Recognition*, Los Angeles, 2004.

[7] M. Gövercin, J. Spehr, E. Steinhagen-Thiessen and F. M. Wahl, “Visual Fall Detection in Home Environments,” in *International Conference for Gerotechnology*, 2008.

[8] S. Wang, S. Zabir and B. Leibe, “Lying pose recognition for elderly fall detection,” in *Robotics: Science and Systems VII*, 2012.