Anderson et al.

D. Anderson, R. Luke, M. Skubic, J.M. Keller, M. Rantz, M. Aud. Evaluation of a video based fall recognition system for elders using voxel space. Gerontechnology 2008; 7(2):68. We are developing a method for recognizing falls through the monitoring of activity from video¹. Privacy is preserved by not using the raw video, but extracting binary silhouette maps, representing the pixels a person occupies in an image. Focus groups at the Aging-in-Place facility of residential apartments known as TigerPlace² indicate that elderly residents are willing to consider silhouette-based images for abnormal event detection such as falls³. In this paper, we present three measures for system performance evaluation and discuss the difficulties in video based human activity recognition of falls and activities representing potential false alarms. Methods An apartment is monitored using multiple cameras and the resulting silhouettes are used to reconstruct a three dimensional representation of the human, called voxel person. Voxel person is tracked and fuzzy logic is used to reason about different types of falls, using a rule base that is validated by our nursing team and designed to recognize typical elderly falls. Results and discussion Activities were performed in an engineering laboratory set up like a living room. Approximately 20 minutes of video, containing 14 falls performed in different ways, are manually analyzed and activities are hand annotated, which is the ground truth for evaluating the automated system. Further data collection will use stunt actors who have been trained by our nursing collaborators to fall in realistic elderly falls. The performance of our fall detection system is evaluated in three ways: (i) matching between the number of linguistic summarizations of human activity, which are automatically extracted from voxel person, and the hand annotated data, (ii) frame-by-frame classification for evaluating how much of each event the system detected, which is important for the recognition of falls in a timely fashion, and (iii) matching between when the subject was on the ground and nurses indicated that a fall should be recognized and what the automated system decided regarding a fall. Confusion matrices are reported for the first two evaluation areas. Recognition rates above 90% were achieved on average for the activities: fallen, sitting on the chair, and lying on the couch. Standing, walking, and kneeling turned out to be somewhat subjective, but were successfully not classified as a fall. All but one of the 14 falls were successfully classified with no false alarms.

References

1. Anderson D, Luke R, Keller J, Skubic M, Rantz M, Aud M. Computer Vision and Image Understanding; under review 2. Rantz M, Porter R, Cheshier D, Otto D, Servey C, Johnson R, Skubic M, Tyrer H, He Z, Demiris G, Lee J, Alexander G, Taylor G. Journal of Housing for the Elderly, in press

3. Demiris G, Rantz M, Aud M, Marek K, Tyrer H, Skubic M, and Hussam A. Medical Informatics and the Internet in Medicine 2004;29, 87-94

Keywords: fall recognition, video, human activity, video privacy, voxel space, voxel person

Address: University of Missouri, USA; E: dtaxtd@mizzou.edu



Figure 1 Two cameras monitoring a scene, (a) voxel person shown sitting in a chair, (b) raw video, and (c and d) the silhouettes used to create voxel person. The large white area in (a) is voxel person, camera locations and view rays are shown, and static surfaces in three dimensional space have their pixel color projected from a camera.