Activity Recognition from Video in Day or Night Using Fuzzy Clustering Techniques

T.Udhayakumar, S.Anandha Saravanan

M.E (Computer Science and Engineering) (student), Dr.Nallini Institute of Engg. & Tech, India

Asst. Professor, Dept of CSE, Dr.Nallini Institute of Engg. & Tech, India

ABSTRACT: An approach for activity state recognition implemented on data collected from various sensors – standard web cameras under normal illumination, web cameras using infrared lighting, and the inexpensive Microsoft Kinect camera system. Sensors such as the Kinect ensure that activity segmentation is possible during day time as well as night. It is especially useful for activity monitoring of older adults since falls are more prevalent at night than during the day. The project is an application of fuzzy set techniques to a new domain. The approach described herein is capable of accurately detecting several different activity states related to fall detection and fall risk assessment including sitting, being upright, and being on the floor to ensure that elderly residents get the help they need quickly in case of emergencies and ultimately to help prevent such emergencies. All detection and fall risk assessment are major goals of research and continue to conduct experiments, in particular, an aging in place facility for the elderly. It describes the silhouette extraction process, the image features employed, and the fuzzy clustering technique used in the work.

KEYWORDS—fuzzy clustering, infrared images, image moments, activity labeling, depth images

I. INTRODUCTION

The aim of the research is to conduct experiments in particular place to provide facility for the elderly and perform detection and fall risk assessment ar. Building a system to discretely monitor the activity of older adults in the apartments, while addressing the privacy concerns, as well as identifying diagnostic measures that are predictors of fall risk, which would then fulfill the long-term goal of the project, to generate alerts that notify caregivers of changes in a resident's condition so they can intervene and prevent or delay adverse health related events.

II. OVERVIEW

Among previous work related to activity analysis, sit-to-stand image descriptors from silhouettes and centroid locations from 3 different camera views. Finally the following features such as the distance of the torso from the feet, the angle created by the torso, head and feet as well as the raw position of the feet and usage of decision tree to identify the activities of sitting and standing. The ground truth used was obtained by hand labeling the transition data from the video sequence for the two individuals tested. Previously simple image subtraction techniques are used to extract the pixels indicating motion and then computed the mean and standard deviation along the horizontal and vertical axes to identify sit, attempt to sit, stand and walk in a given sequence. Checking for motion in the frames in the last two minutes before the stand frame gave the start frame for the beginning of the sit-to-stand activity. The sit-to-stand by finding the points of flexion in the shoulder, knee and hip region using the Gauss-Laguerre Transform, subsequently tracking these natural markers to obtain the trajectories of these points. The angles obtained at the points of flexion were compared to analyze the sit-to-stand activity. While the algorithms mentioned were based on vision sensors under normal illumination. Several experiments indicating the severe fall risk of older adults in low lighting conditions. It created a potential problem since nocturnal activities are an important aspect of an independent lifestyle. This in turn creates a need for surveillance techniques which can be implemented in the absence of light or under negligible lighting conditions.

A study involving dynamic infrared sensors is first implemented in background subtraction, then classified different objects, and finally tracked these objects. The tracker employed iterative systems of location predication and correction based on the location of detected objects in the current frame. To compensate for global motion, , a multi-resolution scheme based on the affine motion model for detecting independent moving objects using forward looking infrared cameras.

The goal of activity recognition is to identify activities as they occur based on data collected by sensors. There exist a number of approaches to activity recognition that vary depending on the underlying sensor technologies that are used to monitor activities, the machine learning algorithms that are used to model the activities and the realism of the testing environment. It postulate that there will exist differences in the interpretation of the activity and thus in the sensor data sequences that correspond to even similar activity labels collected at different sites and annotated by different experimenters.

The labels that are used are provided with each data set as ground truth to train and test our activity recognition algorithms. These charts highlight the fact that there can exist significant differences in the interpretation of an activity name as well as differences in how the activity may be performed at different locations. In order to create activity models that generalize to larger populations, these differences will need to be recognized and standards for activities will need to be defined.

In another approach for activity segmentation, clustering the RGB values of pixels to detect background changes, but the activities were identified using a large data base with prototypes of all the activities. This made the segmentation more supervised in nature. It is also analyzed sit-to-stand by finding the points of flexion in the shoulder, knee and hip region. The RGB values of pixels to detect background changes, but the activities were identified using a large data base with prototypes of all the activities. This made the segmentation more supervised in nature. The application has shown its ability to determine different activities under different environments, both controlled as well as unstructured, and it runs in real time mode.

It can speed up the activity state identification process and lead to faster fall detection. We are also working on sensor fusion algorithms using acoustic sensors to gain more activity related information in the apartments, especially at regions outside the field of view of the sensors.
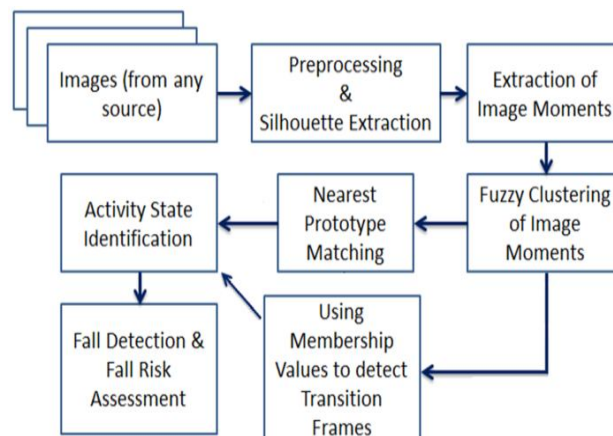

Fig.1.1. Block Diagram of Algorithm of Extracting Image Moments

III. DESIGN CONSIDERATION

Image Extraction

Silhouette extraction and a description of the moments used for clustering describes the fuzzy clustering techniques used for activity analysis and the experimental setup up and results using the standard web cameras under normal illumination. The work of using the web cameras with infrared filters and the Kinect sensor analysis. After obtaining the silhouettes from the image sequence, the next step in the algorithm is extracting image moments as shown in the block diagram in Figure 1.

Image Moments

Image moments are applicable in a wide range of applications such as pattern recognition and image encoding. One of the most widely used is the set of Hu Moments. These are a set of seven central moments taken around the weighted image center. A second set of moment invariants are due to Zernike. The clustering algorithm with applications in several fields such as image processing, pattern recognition, system identification and classification. One of the reasons to choose the GK clustering technique is that it is well suited for the ellipsoidal clusters produced by the Zernike moments. The algorithm is an extension of the Fuzzy C-Means algorithm in which each cluster has its own unique covariance matrix which makes it more robust and more applicable to various data sets which contain ellipsoidal clusters of different orientations and sizes

Gustafson Kessel Clustering

The silhouette extraction process, the image features employed, and the fuzzy clustering technique used in our work. For all three of the image sensors tested in the work of standard web camera under visible lighting, web cameras with IR illumination, and the Kinect sensors, and implemented the same three Zernike image descriptors.

Classifying the transition frames

The results of clustering the Zernike moments from an upright to sit sequence are noted. The two clusters are color coded with red indicating the sit activity and blue indicating upright. While the clustering technique separates the two activity frames, the labeling itself is implemented by the semi supervised approach.

IV. IMPLEMENTATION AND RESULT

A successful and yet simple technique for detecting activity frames using fuzzy clustering methods. A soft classifier was constructed from the clustering results, and activity classification results using the Zernike Moments were obtained. The fuzzy clustering technique was compared against the Vicon Motion Capture system in the laboratory and was shown to be accurate in activity recognition.

The technique has shown the ability to link the results of one sequence to the soft classification of others, showing its strength and success over a range of image sensors and in real living quarters. Preliminary experiments were conducted to establish the input parameters and best features to use for this data. Several participants performed different activities. Silhouettes were extracted from the raw image sequences, and the moment features were computed. The GK clustering technique requires the number of clusters to be specified as an input parameter.

In preliminary experiments, clustering the Zernike moments using the GK algorithm with the number of clusters initialized to the number of activities yielded the best results. Since single camera images are used here, the activities of walking and standing cannot be differentiated in general, they are grouped together as upright frames for

the purpose of activity recognition. The background model is initialized, regions in subsequent images with significantly different characteristics from the background are considered as foreground objects. Areas classified as background are also used to update the background model.
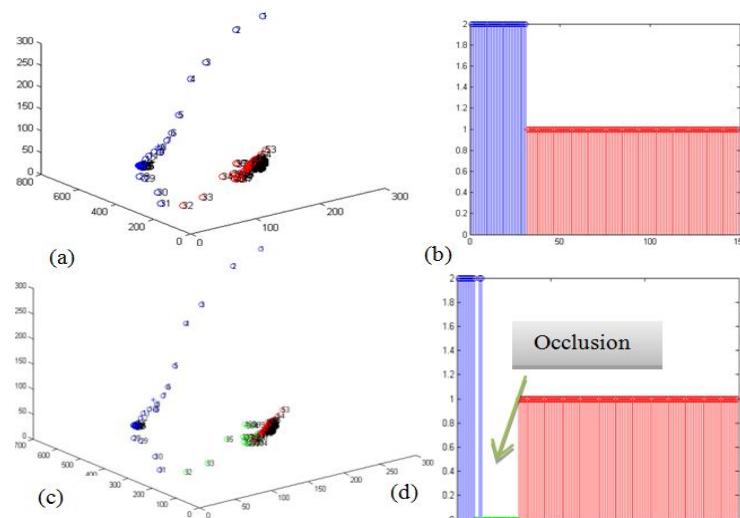

Fig 4.1 Test results on a depth video sequence with two activities – sit and upright.

V.   CONCLUSION

It is a successful and yet simple technique for detecting activity frames using fuzzy clustering methods. A soft classifier was constructed from the clustering results, and activity classification results using the Zernike Moments were obtained. The fuzzy clustering technique was compared against the Vicon Motion Capture system in the laboratory and was shown to be accurate in activity recognition. The technique has shown the ability to link the results of one sequence to the soft classification of others, showing its strength and success over a range of image sensors and in real living quarters.

Future work includes creating an online version of the fuzzy clustering algorithms described above and validating the results using fuzzy validity measures. This can speed up the activity state identification process and lead to faster fall detection. Apart from that, future work also includes finding ways to detect occlusion in image silhouettes. Experiments are currently being conducted to test the algorithm under different scenarios with different activities to test the performance of the algorithm for different locations of the Kinect sensor.

REFERENCES

[1]   Allin.S and Mihailidis.A (2008), "Automated sit-to-stand detection and analysis." AAAI 2008 Fall Symposium on AI in Eldercare: New Solutions to Old Problems, Arlington, VA.

[2]   Babuska.R, van der Veen.P.J, Kaymak.U, (2002) "Improved covariance estimation for gustafson-kessel clustering," IEEE International Conference on Fuzzy Systems pp. 1081–1085.

[3]   Banerjee.T, Keller.JM, Skubic.M & Abbott.CC (2010), "Sit-To-Stand Detection Using Fuzzy Clustering Techniques," IEEE International Conference on Fuzzy Systems (FUZZ-IEEE) at the World Conference on Computational Intelligence (WCCI), Spain.

[4]   Berg.K, S. Wood-Dauphinee, Williams J.I, Maki B," (1992) Measuring balance in the elderly: Validation of an instrument". Can. J. Pub. Health 2:S7-11.

[5]   Fang.H and Lei.H (2012), "BP Neural Network for Human Activity Recognition in Smart Home." Computer Science & Service System (CSSS), International Conference on. IEEE.

[6]   Foroughi.H, Aski.B.S, and Pourreza.H (2008), "Intelligent video surveillance for monitoring fall detection of elderly in home environments," in Proc. 11th International Conference on Computer and Information Technology, Khulna, Bangladesh, pp. 219-224.

[7]   Hu.M.K, "Visual pattern recognition by moment invariants", (1962) IRE Trans. on Information Theory, IT-8:pp. 179-187, 1962.

[8]   Iosifidis.A, Tefas.A, Pitas.I, "Eating and drinking activity recognition based on discriminant analysis of fuzzy distances and activity volumes," Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on , pp.2201-2204, 25-30 March 2012.

[9]   Maadi.A.EI and Maldague.X (2007), Outdoor infrared video surveillance "A novel dynamic technique for the subtraction of a changing background of IR images." Infrared Physics & Technology, 49(3)p p. 261-265.

[10]   Oikonomopoulos.A, Pantic.M, and IPatra.I (2009),"Sparse B spline polynomial descriptors for human activity recognition," Image and Vision Computing, 27(12), pp. 1814-1825.

[11]   Stauffer.C and Grimson.W.E.L (2000), "Learning patterns of activity using real-time tracking," IEEE Trans. Pattern Anal. Mach. Intell., vol. 22, pp. 747–757.

[12]   Sung.J, Ponce.C, Selman.B, and Saxena.A (2011), "Human activity detection from rgbd images", In AAAI workshop on Pattern, Activity and Intent Recognition.

[13]   Urso.P.D, Massari.R, (2013) "Fuzzy clustering of human activity patterns," Fuzzy Sets and Systems", vol. 215, pp. 29-54.

[14]   Xue, Ming.Z, Song.D, Wan.W, Jin.B, Infrared gait.S, (2010) "Recognition based on wavelet transform and support vector machine. Patt. Recog", 43, Vol. 8, 2904–2910.