International Journal For Technological Research In Engineering
Volume 6, Issue 5, January-2019 *ISSN (Online): 2347 - 4718*

AUTOMATIC MOTION BASED PERSON ACTIVITY RECOGNITION IN VIDEO SURVEILLANCES

C.Arul
Senior Lecturer, Dept. of ECE,
Bhaktavatsalam Polytechnic college, Karaipettai, Kancheepuram, Tamil Nadu, India

ABSTRACT: Computer vision includes methods for acquiring, processing, analysing, and understanding images. Applications of this field include detecting events, con- trolling processes, navigation, modelling objects or environments, automatic inspection and many more. Activity recognition is one of the applications of computer vision that aims to recognize the actions and goals of one or more agents from a series of observations on the agent's actions and the environmental conditions. The goal of the project is to train an algorithm to automate detection and recognition of human activities performed in the video data. The project can be utilized in scenarios such as surveillance systems, intelligent environment, sports play analysis and web based video retrieval. In the proposed system, actions like walking, running, bending, and side galloping and hand wave were recognized using a Temporal Template matching technique called Motion History Image (MHI) methodology. MHI method is extensively being used to represent the history of temporal changes involved in the execution of an activity. Here the intensity of the pixels of a scalar valued image is varied, depending upon the motion history.
Keywords: Activity Recognition, MHI, Human Detection.

I. INTRODUCTION

Human action recognition is the process of labelling video sequences containing human action with corresponding action classes. More specifically, vision-based human action recognition is discussed and studied in this report. Vision-based human action recognition is the process of recognizing the human actions in video sequences by utilizing computer vision techniques. Due to the increase of digital video cameras used in everyday life, more and more video content is generated and uploaded to the Internet or stored in large video dataset. Categorizing rich video content based on the actions appearing in the video is a good way to reach the initial goal of organizing these videos. Also, human action recognition is a popular research area due to its potential application in visual surveillance; content based video retrieval, human-computer interaction and sports annotation. For example, with successful human action recognition, the visual surveillance system in large public area can automatically extract high-level semantic information from the surveillance video thus making it possible to make alarms to the public when predefined dangerous behaviours occur in the range of surveillance; content-based video retrieval system can search and locate the video content with specific definition fast and precisely; human-computer interaction systems can smoothly interact with human body movement and provide more sophisticated interface.

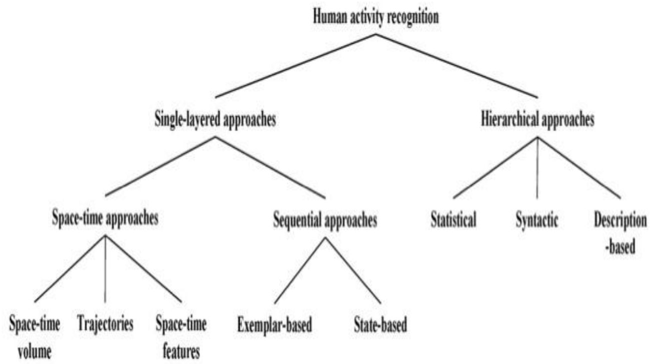

Figure 1: Approaches for Activity Recognition

II. RELATED WORKS

Computer vision is a field that includes methods for acquiring, processing, analyzing, and understanding images and, in general, high-dimensional data from the real world in order to produce numerical or symbolic information, e.g., in the forms of decisions [1]. Activity recognition is one of the sub-fields of computer vision that aims to recognize the actions and goals of one or more agents from a series of observations on the agents' actions and the environmental conditions [2]. Human activity recognition focusses on accurate detection of human activities based on a pre-defined activity model. It can be exploited to great societal benefits, especially in real life, human-centric applications such as eldercare and healthcare [3].

III. EXISTING METHODOLOGY

High-level action units to represent human actions in videos and, based on such units, a novel sparse model is developed for human action recognition. There are three interconnected components in our approach. First, we propose a new context-aware spatial temporal descriptor, named locally weighted word context, to improve the discriminability of the traditionally used local spatial-temporal descriptors. Second, from the statistics of the context-aware descriptors, we learn action units using the graph regularized nonnegative matrix factorization, which leads to a part-based representation and encodes the geometrical information. These units effectively bridge the semantic gap in action recognition. Third, we propose a sparse model based on a joint 1-norm to preserve the representative items and suppress noise in the action units. Intuitively, when learning the dictionary for action representation, the sparse model captures the fact that actions from the same class share similar units. The proposed approach is evaluated on several publicly available data sets. The experimental results and analysis clearly demonstrate the effectiveness of the proposed approach.

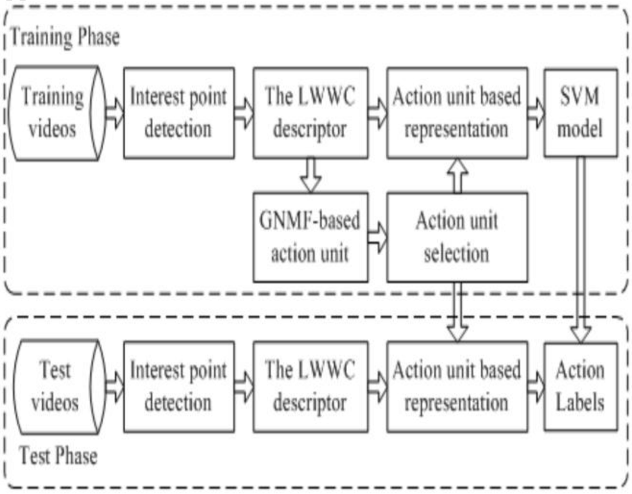

Figure 2 Flowchart For Action Recognition

IV. PROPOSED METHODOLOGY

The interactions among people are important cues for recognizing collective activity. In this paper, we concentrate on modelling the person–person interactions for collective activity recognition. Rather than relying on hand-craft description of the person–person interaction, we propose a novel learning-based approach that is capable of computing the class-specific person–person interaction patterns. In particular, we model each class of collective activity by an interaction matrix, which is designed to measure the connection between any pair of atomic activities in a collective activity instance. We then formulate an interaction response (IR) model by assembling all these measurements and make the IR class specific and distinct from each other. A multitask IR is further proposed to jointly learn different person–person interaction patterns simultaneously in order to learn the relation between different person–person interactions and keep more distinct activity specific factor for each interaction at the same time. Our model is able to exploit discriminative low-rank representation of person–person interaction. Experimental results on two challenging data sets demonstrate our proposed model is comparable with the state-of-the-art models and show that learning person–person interactions plays a critical role in collective activity recognition.

Due to this novel approach, our method is robust to pose and alignment and hence can be used to recognize faces from unconstrained videos successfully. Moreover, in a moving scene, camera angle, illumination and other imaging conditions may change quickly leading to performance loss in accuracy. In such situations, it is impractical to re-enrol the individual and re-train the classifiers on a continuous basis. Our novel approach addresses these practical issues. An experimental result on the well-known YouTube Face database demonstrates the effectiveness of our method.

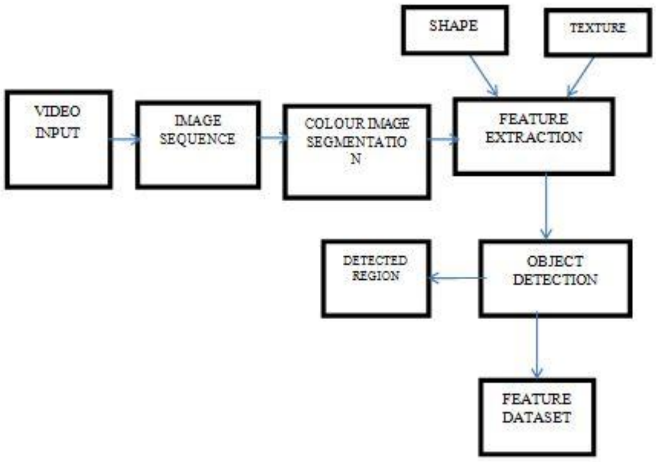

Figure 3 Block Diagram of Proposed System

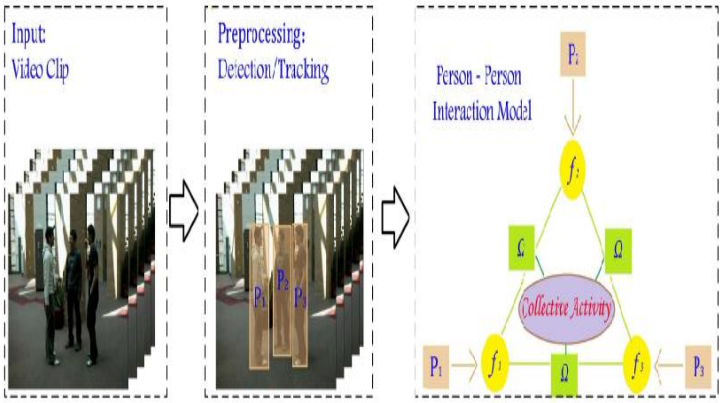

Fig 4.Interaction Model

MODULES

1. Input video
2. Converted into frames
3. Feature Extraction
4. Person Activity Recognition
5. Tracking

Input video

The video input is live streaming video or stored video in any format like .mp4, .mpeg, or .avi. This Live video or recorded video can be given to MALTAB; with the help of Computer Vision system toolbox this video can be further processed.

Video Converted to Sequence of Image

Frames can be obtained from a video and converted into images. To convert a video frame into an image some subroutine function is used

Feature extraction

Feature extraction is based on Shape and Texture. The shape tells us the size of the features of the object , the texture is used to find some statics data for features of the object.

Person Activity Detection

Person activity detection is that the method of finding instances of real-world entities likes faces, bicycles, and buildings in pictures or videos. Person Activity algorithms usually use the extracted options associated with the learning algorithms to acknowledge the instances of an object class. It is widely employed in applications like image retrieval, security, police investigation, and automatic vehicle parking systems.

Tracking
The tracking helps us to find the individual or group person monitoring.

V. IMPLEMENTATION RESULTS

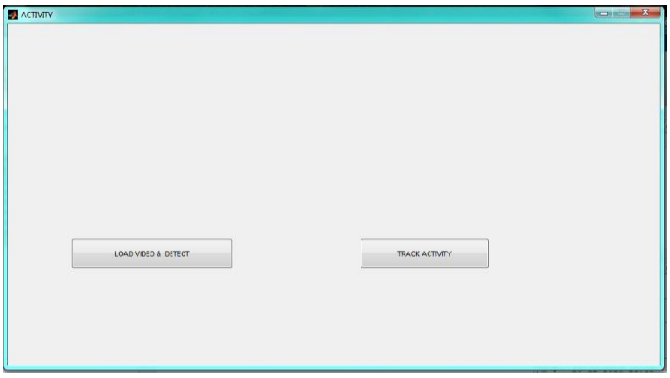

Figure 4 shows the input video is loaded and detect that video then track the activity of person.

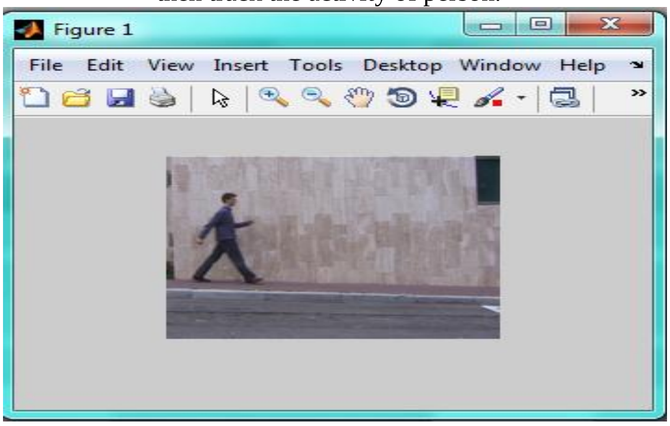

Figure 6 shows the converted frame 1 from the loaded video.

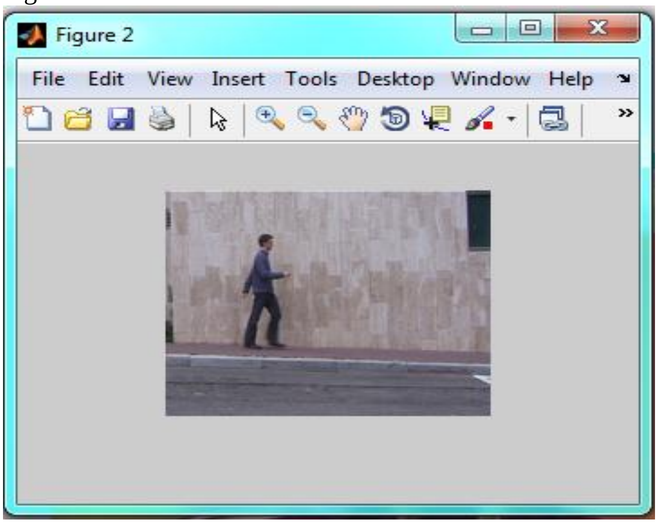

Figure 7 shows the converted frame 2 from the loaded video.

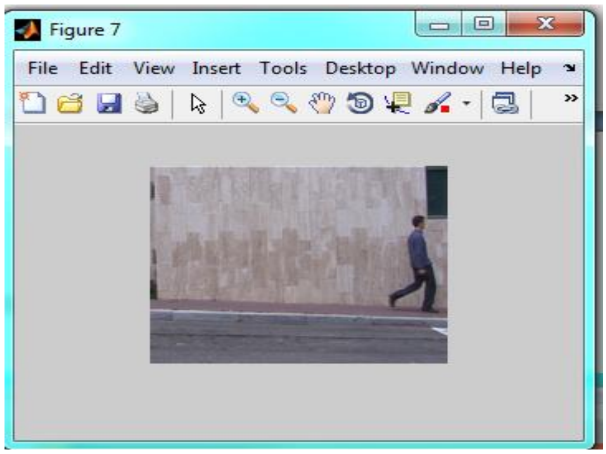

Figure 8 shows the converted frame 3 from the loaded video.

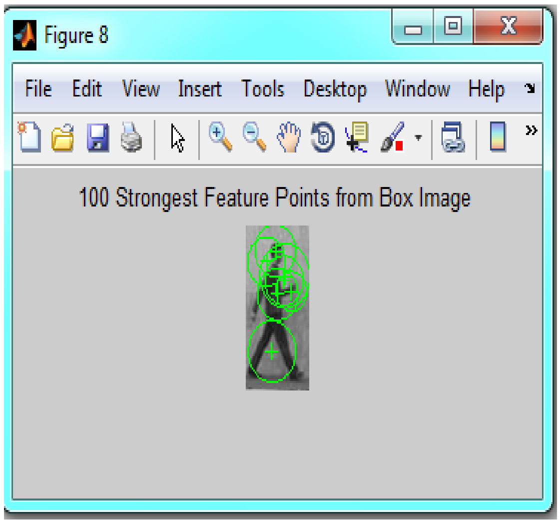

Figure 9 shows the feature extraction of the walking frame.

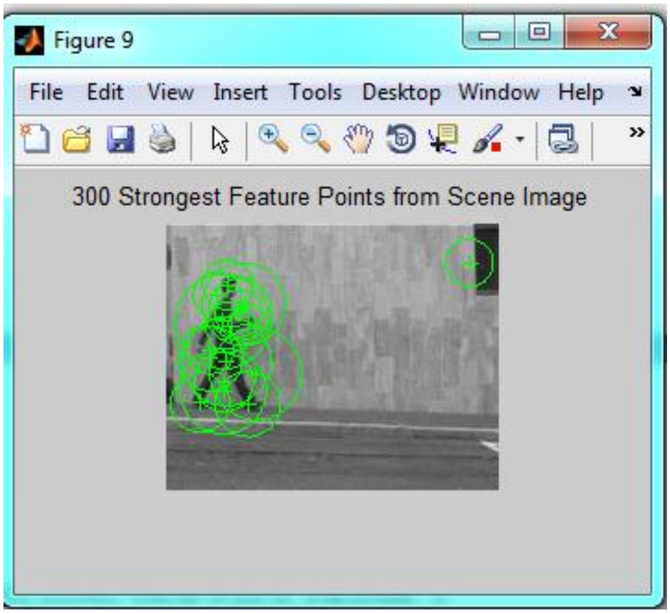

Figure 10 shows the feature extraction in video frame

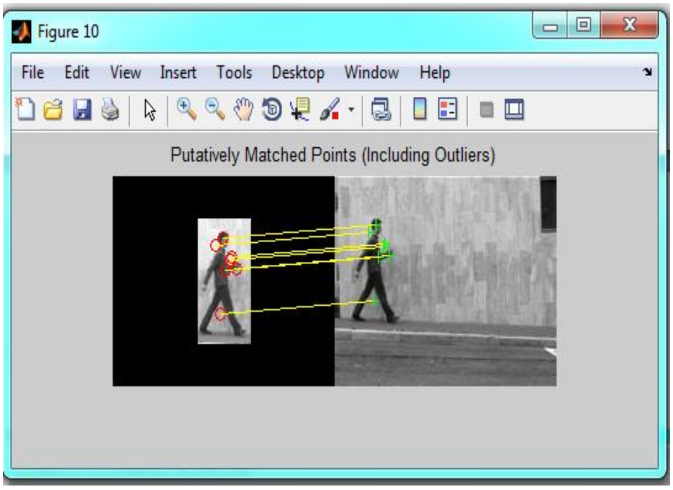

Figure 11 shows the feature matches with training and test video frames

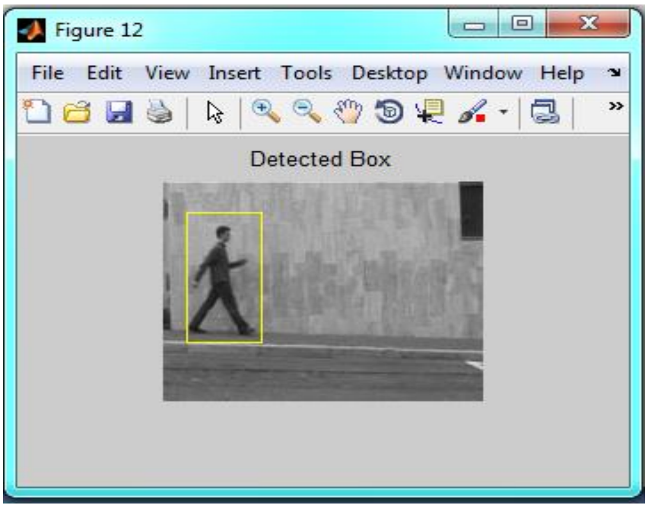

Figure 12 shows the person detected region

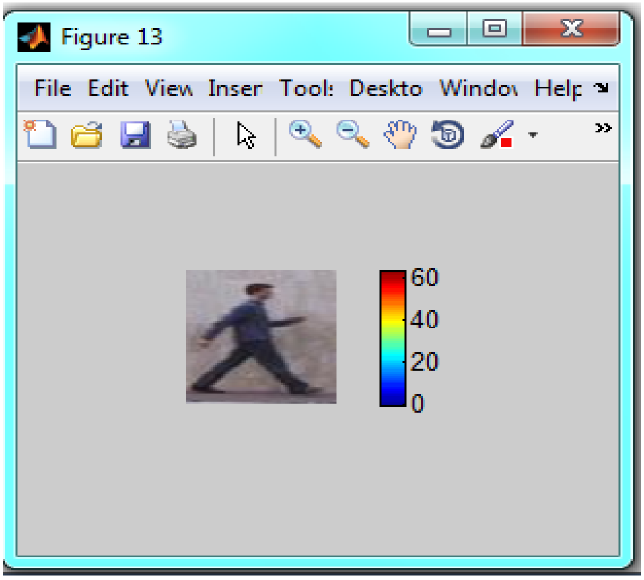

Figure 13 shows the colour descriptor range

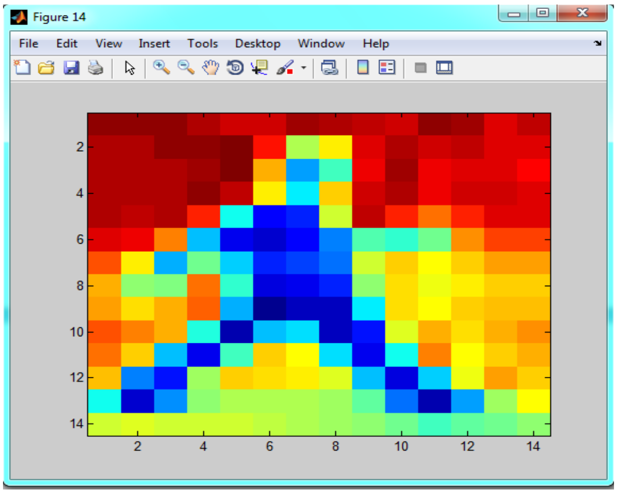

Figure 14 shows walking colour pixels

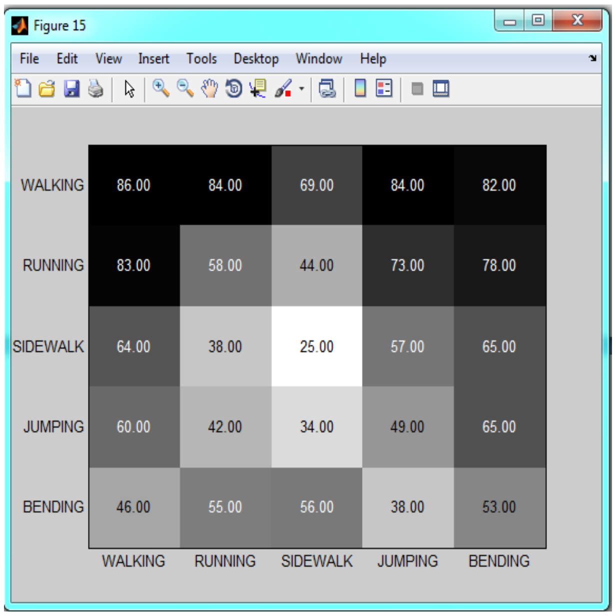

Figure 15 shows the confusion matrix

VI. CONCLUSION

This decision is made based on early anticipations of the experiments that we will need later in order to dissect influence of scene context in action recognition .We would like to investigate the effect of motion/shape/color properties on the joint distribution of action and context. One of the conclusion made in proposed system was that they showed that SIFT features basically captures more of the static appearance of the video scene while hog of features captures dynamic information which is basically taken from the moving objects.

REFERENCES

[1] J. Aggarwal and M. Ryoo, "Human activity analysis: A review," ACM Computing Surveys (CSUR), vol. 43, pp. 16:1–16:43, april 2011.

[2] A. Yilmaz and M. Shah, "Recognizing human actions in videos acquired by uncalibrated moving cameras," in Proceedings of the Tenth IEEE In-

ternational Conference on Computer Vision (ICCV'05), vol. 1. IEEE Computer Society, 2005, pp. 150–157.

[3] E. Shechtman and M. Irani, "Space-time behavior based correlation," in IEEE Conference on Computer Vision & Pattern Recognition (CVPR), vol. 1. IEEE Computer Society, June 2005, pp. 405–412.

[4] S. M. Khan and M. Shah, "Detecting group activities using rigidity of for- mation," in Proceedings of the 13th annual ACM international conference on Multimedia, 2005, pp. 403–406.

[5] N. Nguyen, D. Phung, S. Venkatesh, and H. Bui, "Learning and detecting activities from movement trajectories using the hierarchical hidden markov model," in IEEE Computer Society Conference on Computer Vi- sion & Pattern Recognition CVPR, vol. 2, Jan. 2005, pp. 955–960.

[6] S. W. Joo and R. Chellappa, "Attribute grammar-based event recogni- tion and anomaly detection," in IEEE Computer Society Conference on Computer Vision & Pattern Recognition Workshop CVPRW, Jan. 2006, pp. 107–112.

[7] A. Gupta, P. Srinivasan, J. Shi, and L. Davis, "Understanding videos, constructing plots learning a visually grounded storyline model from an- notated videos," in IEEE Computer Society Conference on Computer Vi- sion & Pattern Recognition CVPR, Jan. 2009, pp. 2012–2019.

[8] Y. Ke, R. Sukthankar, and M. Hebert, "Spatio-temporal shape and flow correlation for action recognition," in IEEE Computer Society Conference on Computer Vision & Pattern Recognition (CVPR), June 2007, pp. 1–8.

