Analysis of Gait Pattern to Recognize the Human Activities

Jay Prakash Gupta[1], Pushkar Dixit[2], Nishant Singh[3], Vijay Bhaskar Aemwal[4]

[1]Infosys Limited, Pune, India
[2]Faculty of Engineering and Technology Agra College, Agra, India
[3]Poornima Institute of Engineering and Technology, Jaipur, India
[4]SiemensInformation System, India

Abstract – **Human activity recognition based on the computer vision is the process of labelling image sequences with action labels. Accurate systems for this problem are applied in areas such as visual surveillance, human computer interaction and video retrieval. The challenges are due to variations in motion, recording settings and gait differences. Here we propose an approach to recognize the human activities through gait. Activity recognition through Gait is the process of identifying an activity by the manner in which they walk. The identification of human activities in a video, such as a person is walking, running, jumping, jogging etc are important activities in video surveillance. We contribute the use of Model based approach for activity recognition with the help of movement of legs only. Experimental results suggest that our method are able to recognize the human activities with a good accuracy rate and robust to shadows present in the videos.**

Keywords – **Feature Extraction, Gait Pattern, Human Computer Interaction, Activity Recognition, Video Surveillance**

I. Introduction

THE goal of automatic video analysis is to use computer algorithms to automatically extract information from unstructured data such as video frames and generate structured description of objects and events that are present in the scene. Among many objects under consideration, humans are of special significance because they play a major role in most activities of interest in daily life. Therefore, being able to recognize basic human actions in an indispensable component towards this goal and has many important applications. For example, detection of unusual actions such as jumping, running can provide timely alarm for enhanced security (e.g. in a video surveillance environment) and safety (e.g. in a life-critical environment such as a patient monitoring system). In this paper, we use the concept of Gait for human activity recognition. The definition of Gait is defined as: "A particular way or manner of moving on foot". Using gait as a biometric is a relatively new area of study, within the realms of computer vision. It has been receiving growing interest within the computer vision community and a number of gait metrics have been developed. We use the term Gait recognition to signify the identification of an individual from a video sequence of the subject walking. This does not mean that Gait is limited to walking, it can also be applied to running or any means of movement on foot. Gait as a biometric can be seen as advantageous over other forms of biometric identification techniques for the following reasons: unobtrusive, distance recognition, reduced detail, and difficult to conceal. This paper focuses on the design, implementation, and evaluation of activity recognition system through gait in video sequences. It introduces a novel method of identifying activities only on the basis of leg components and waist component. The use of waist below components for recognizing the activities makes it to achieve fast activity recognition over the large databases of videos and hence improves the efficiency and decreases the complexity of the system. To recognize the actions, we establish the features of each action from the parameters of human model. Our aim is to develop a human activity recognition system that must work automatically without human intervention. We recognized four actions in this paper namely walking, jumping, jogging and running. The walking activity is identified by the velocities of all components superior to zero but lesser than a predefined threshold. In case of jumping activity, every part of human moves only vertically and in the same direction either up or down. Therefore, jumping action can be identified by the velocities of all the three components to be near or equal to zero in horizontal direction but greater than zero in vertical direction. The only differences between jogging and running activities are that travelling speed of running is greater than jogging and other difference is of distance ratio between the leg components to the axis of ground. In case of running activity, speed of travelling is greater than jogging and the other difference is of distance ratio between leg components to the axis of ground.

The rest of the paper is structured as follows: Section 2 discusses the trend of activity recognition research area in the past decade which introduces the fundamentals of gait recognition systems and human activity recognition models; Section 3 presents the proposed work of human activity recognition using Gait; Section 4 analyzes and evaluates the empirical results of experiments to validate the proposed framework. Before evaluating the proposed system, some hypotheses are established and the evaluations are conducted

against these hypotheses; finally section 5 summarizes the novelties, achievements, and limitations of the framework, and proposes some future directions of this research.

II. LITERATURE REVIEW

In recent years, various approaches have been proposed for human motion understanding. These approaches generally fall under two major categories: model-based approaches and model-free approaches. Poppe has made a survey on vision based human action recognition [1]. When people observe human walking patterns, they not only observe the global motion properties, but also interpret the structure of the human body and detect the motion patterns of local body parts. The structure of the human body is generally interpreted based on their prior knowledge. Model-based gait recognition approaches focus on recovering a structural model of human motion, and the gait patterns are then generated from the model parameters for recognition. Model-free approaches make no attempt to recover a structural model of human motion. The features used for gait representation includes: moments of shape, height and stride/width, and other image/shape templates.

Leung & Yang reported progress on the general problem of segmenting, tracking, and labeling of body parts from a silhouette of the human [2]. Their basic body model consists of five U-shaped ribbons and a body trunk, various joint and mid points, plus a number of structural constraints, such as support. In addition to the basic 2-D model, view-based knowledge is defined for a number of generic human postures (e.g., "side view kneeling model," "side horse motion"), to aid the interpretation process. The segmentation of the human silhouette is done by detecting moving edges. Yoo et al. estimate hip and knee angles from the body contour by linear regression analysis [3]. Then trigonometric-polynomial interpolant functions are fitted to the angle sequences and the parameters so-obtained are used for recognition.

In [4], human silhouette is divided into local regions corresponding to different human body parts, and ellipses are fitted to each region to represent the human structure. Spatial and spectral features are extracted from these local regions for recognition and classification. In model-based approaches, the accuracy of human model reconstruction strongly depends on the quality of the extracted human silhouette. In the presence of noise, the estimated parameters may not be reliable.

To obtain more reliable estimates, Tanawongsuwan and Bobick reconstruct the human structure by tracking 3D sensors attached on fixed joint positions [5]. However, their approach needs lots of human interaction because they have considered and identified only walking type of activity whereas our method has considered four type of activities and the performance is reasonable for each type of activity. Wang et al. build a 2D human cone model, track the walker under the Condensation framework, and extract static and dynamic features from different body part for gait recognition [6]. Their approach has fused static and dynamic features to improve the gait recognition accuracy but extraction of both static and dynamic features required more computation which lacks its applicability in real time scenario.

Zhang et al. used a simplified five-link biped locomotion human model for gait recognition [7]. Gait features are first extracted from image sequences, and are then used to train hidden Markov models for recognition. In [8], an approach for automatic human action recognition is introduced by using the parametric model of human from image sequences using motion/texture based human detection and tracking. They used the motion/texture of full body part whereas proposed approach used only the gait pattern of the lower body part which is more time efficient. Bobick & Davis interpret human motion in an image sequence by using *motion-energy* images (MEI) and *motion-history* images (MHI) [9]. The motion images in a sequence are calculated via differencing between successive frames and then thresholded into binary values. These motion images are accumulated in time and form MEI, which are binary images containing motion blobs. The MEI is later enhanced into MHI, where each pixel value is proportional to the duration of motion at that position. Moment-based features are extracted from MEIs and MHIs and employed for recognition using template matching. Because this method is based on the whole template matching instead of the only gait pattern of the legs, it does not take the advantage of recent development whereas we incorporated the matching only based on the gait analysis. Recent Gait studies for activity recognition suggest that gait is a unique personal characteristic, with cadence and cyclic in nature [10]. Rajagopalan & Chellappa [11] described a higher-order spectral analysis-based approach for detecting people by recognizing human motion such as walking or running. In their proposed method, the stride length was determined in every frame as the image sequence evolves.

TABLE 1. COMPARISON OF EXISTING APPROACHES

Ref.	Method	Advantage	Disadvantage	Uses
[8]	Gait recognition	Locomotion human model	Insensitive to noise	Indoor scenario
[9]	Model-based Action Recognition	Inclusion of motion texture	Poor performance in walking case	Indoor environment
[12]	Spectral analysis of human motion	Higher-order Spectral	Periodic detection	Differentiate between people and vehicular objects
[13]	View based motion analysis	Object models are not required	Need to reduce the distribution combinatory	Outdoor scenario
[27]	Activity recognition using smartphones	Real time application	More than one classifier reduces the accuracy	Indoor/Outdoor both

Vega and Sarkar [12] offered a novel representation scheme for view-based motion analysis using just the change in the relational statistics among the detected image features, without the need for object models, perfect segmentation, or part-level tracking. They modeled the relational statistics using the

probability that a random group of features in an image would exhibit a particular relation. To reduce the representational combinatorics of these relational distributions, they represented them in a Space of Probability Functions (SoPF). Different motion types sweep out different traces in this space. They also demonstrated and evaluated the effectiveness of that representation in the context of recognizing persons from gait. But, there method requires multiple cameras from different viewpoints to model multi-view recognition system which requires extra setup and also computation, whereas the proposed approach is able to achieve high recognition performance from only a single viewpoint. Several other approaches and features used in [13-25] may be tied with gait analysis to predict the human actions. Human activity recognition using smartphones is also studied [26] but its recognition rate can be improved using gait analysis with more time efficiently. Table 1 compares the existing approaches.

III. PROPOSED METHODOLOGY

The proposed technique of human activity recognition is based on the foreground extraction, human tracking, feature extraction and recognition. Figure 1 shows the framework of the introduced human activity recognition system using Gait to identify four basic human activities (i.e. walking, running, jogging and jumping). The proposed method has following main steps: Foreground Extraction, Human Tracking, Feature Extraction and Activity Recognition. In this framework, the video is given as an input to the system from the activity database and frames are extracted from that video. The parametric model of human is extracted from image sequences using motion/texture based human detection and tracking. After that the results are displayed as the recognized activities like walking, running, jogging and jumping; and finally the performance of the method is tested experimentally using the datasets under indoor and outdoor environments.

A. Foreground Extraction

The first step is to provide a video sequence of an activity as an input in the proposed system from the dataset. That video contains a number of continuous frames. After that background subtraction technique is used to separate moving object present inside those frames. But these frames contain some noises which may lead to incurrent foreground subtraction. So first of all, we remove these noises. Some of the small noises are removed by using morphological image processing tools such as Erosion, Dilation, or Gaussian Filters. Generally, an object might be detected in several fragmented image regions. In that case, a region-fusion operation is needed. Two regions are considered to be the same object if they are overlapped or their distance less than a specific threshold value. With these constraints, the method is again very sensible to light condition, such as shadow, contrast changing and sudden changes of brightness.
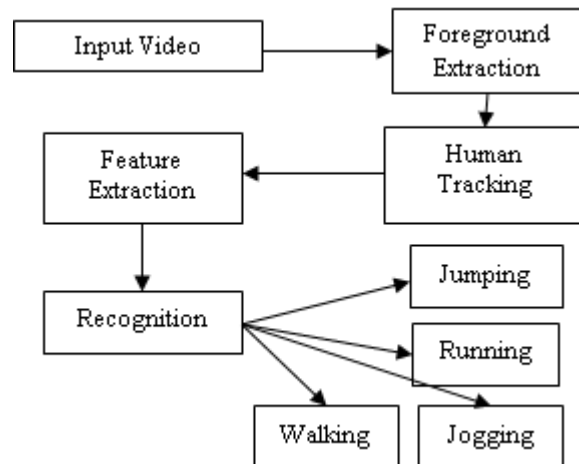

Fig. 1. Framework of Proposed System of Human Activity recognition

Intuitively, introducing some special characteristics of object, for instance texture properties, will probably improve the better results. Therefore, in the fusion process the color probability density of object's texture is additionally applied for computing the similarity between regions using Mean-shift algorithm [27]. This mixture of motion and texture of object for detection and tracking can reduce significantly noises and increases consequently the effectiveness of our tracking algorithm. However, there are always additive noises superposed with detected objects that will be eliminated later by human model constraints. The mean shift algorithm is a nonparametric clustering technique which does not require prior knowledge of the number of clusters, and does not constrain the shape of the clusters. Hence, mean shift represents a general non-parametric mode finding/clustering procedure.

B. Human Tracking and Activity Recognition

In this phase, we apply Hu-moments [28] for shape analysis in which Zero- to third-order moments are used for shape recognition and orientation as well as for the location tracking of the shape. Hu-moments are invariant to translation, rotation and scaling. Hu derived expressions from algebraic invariants applied to the moment generating function under a rotation transformation. They consist of groups of nonlinear centralized moment expressions. The result is a set of absolute orthogonal (i.e. rotation) moment invariants, which can be used for scale, position, and rotation invariant pattern identification. The advantage of using Hu invariant moment is that it can be used for disjoint shapes. In particular, Hu invariant moment set consists of seven values computed by normalizing central moments through order three. In terms of central moment the seven moments are given as below:

$$M_1 = \eta_{20} + \eta_{02}$$

$$M_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2$$

$$M_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2$$

$$M_4 = (h_{30} + h_{12})^2 + (h_{21} + h_{03})^2$$

$$M_5 = (h_{30} - 3h_{12})(h_{30} + h_{12})[(h_{30} + h_{12})^2 - 3(h_{21} + h_{03})^2]$$
$$+ (3h_{21} - h_{03})(h_{21} + h_{03})[3(h_{30} + h_{12})^2 - (h_{21} + h_{03})^2]$$

$$M_6 = (h_{20} - h_{02})(h_{30} + h_{12})^2 - (h_{21} + h_{03})^2$$
$$+ [4h_{11}(h_{30} + h_{12})(h_{21} + h_{03})]$$

$$M_7 = (3h_{21} - h_{03})(h_{30} + h_{12})[(h_{30} + h_{12})^2 - 3(h_{21} + h_{03})^2]$$
$$+ (3h_{21} - h_{30})^2(h_{21} + h_{03})[3(h_{30} + h_{12})^2 - (h_{21} + h_{03})^2]$$

These seven values given by Hu are used as a feature vector for centroid in the human model.

C. Feature Extraction

We employed a model based approach to extract the features. The extracted foreground that supposed to be a human is segmented into centroid and two leg components. We use Mean-shift algorithm again for computing the similar regions below the centroid of the human body for each leg components that will serve for tracking legs. We assume that with only these three components of human model the four basic actions could be identified correctly. The human model constraints are used for noise suppression. The three components namely centroid, left leg and right leg (i.e. vm1, vm2, vm3 respectively), are used in order to model parametric approach. The threshold concept is also used along with the defined method. Threshold calculation is applied as follows: Video sequences from the KTH and Weizmann datasets are normalized on the basis of number of frames and the time of a particular sequence for an activity. The threshold is calculated on the basis of a case study given in [29]. To recognize the actions, we establish the features of each action from the parameters of human model as follows: **Walking feature:** In case of walking action, every part of human move generally and approximately in the same direction and speed. Therefore, the walking activity can then be identified by the velocities of all components superior to zero but lesser than a predefined threshold for walking. Note that the significant difference between running and walking strides is that at least one of the feet will be in contact with the principal axis (ground) at any given time as shown in Figure 2 (a). **Jumping feature:** In case of jumping activity, every part of human moves only vertically and in the same direction either up or down [30-39]. Therefore, jumping action can be identified by the velocities of all the three components to be near or equal to zero in horizontal direction but greater than zero in vertical direction as shown in Figure 2(b). **Jogging feature:** The only differences between jogging and running activities were that travelling speed of running is greater than jogging and other difference is of distance ratio between the leg components to the axis of ground as shown in Figure 2(c). **Running feature:** Similarly in case of running activity, speed of travelling is greater than jogging and the other difference is of distance ratio between leg components to the axis of ground as shown in Figure 2 (d).
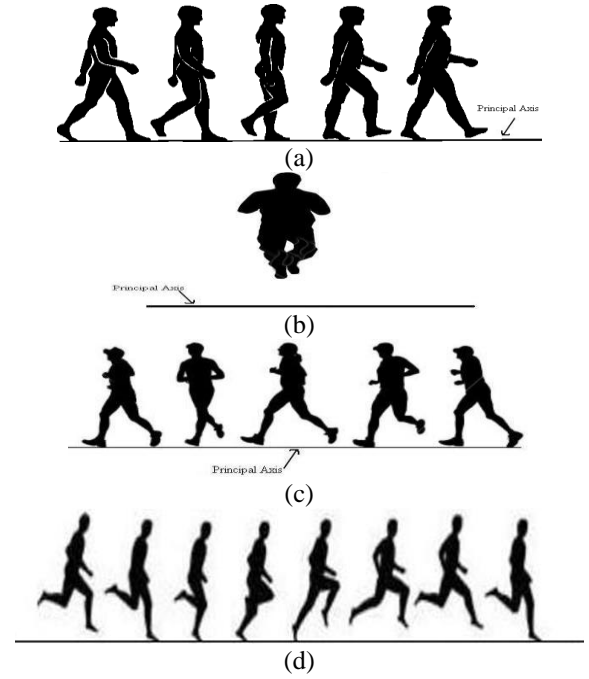

Fig. 2. Silhouette pattern for (a) Walking, (b) Jumping, (c) Jogging and (d) Running

Algorithm for Human Activity Recognition

1) Input is fed to the system as a single video sequence.
2) Frames are extracted from the input video, which are used for further processing.
3) Background subtraction technique is implemented to subtract background from the frames in order to obtain the foreground moving object.
4) Morphological operators are used to remove additional noises in the frames.
5) Mean-shift algorithm is used to track the human; based on the texture similarities in the frames.
6) Hu-moments are calculated to recognize the centroid of the tracked human. Again the Mean-shift algorithm is used to recognize each leg components of the model.
7) For feature extraction, model based approach is employed. The extracted foreground that supposed to be human is then segmented into centroid and the two leg components i.e., total three components.
8) The features of each action from the parameters of human model acts as the features for classifying all four activities (walking, jumping, jogging and running).
9) The features depend on the following criteria: Walking, Jumping, Jogging and Running.

(a)   (b)

(c)   (d)
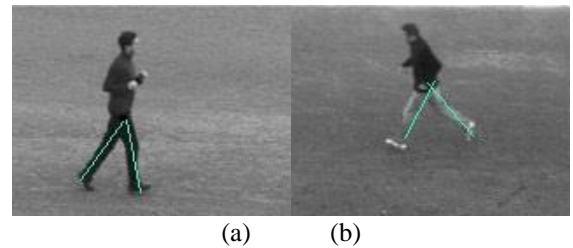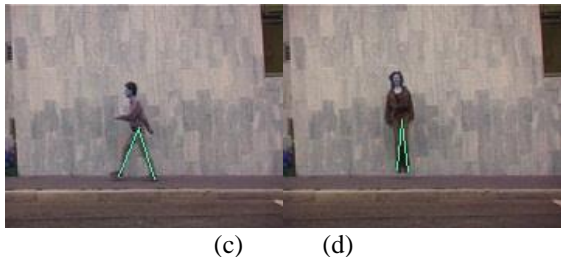
Fig. 3. Templates of (a) jogging, (b) running, (c) walking, and (d) jumping for human activities.

IV.   RESULTS AND DISCUSSIONS

This section analyses the various aspects of the proposed method. In activity recognition through gait, feature requirement is the main issue to model the human according to the parameters to fulfill the criteria.

A. Data Set Used

In order to evaluate our proposed approach of human activity recognition, we have used two datasets: (1) KTH Human Actions dataset (http://www.nada.kth.se/cvap/actions) and (2) Weizmann Actions dataset (http://www.wisdom.weizmann.ac.il/~vision/SpaceTimeActions.html).

KTH Human Actions dataset: KTH video dataset uses six types of human actions such as "walking", "jogging", "running", "boxing", "hand waving" and "hand clapping", which were performed by 25 subjects in different scenarios with different clothing conditions as well.

The video sequences are down sampled to 160*120 pixels and an average length varying from 4 to 41 seconds. This dataset contains 2391 activity sequences. All videos are having static background with 25 fps. We use walking, jogging and running sequences of KTH actions data set for evaluation.

Weizmann Actions dataset: Weizmann Actions dataset uses ten types of natural human actions such as "run," "walk," "skip," "jumping-jack", "jump-forward-on-two-legs", "jump-in-place-on-two-legs", "gallop sideways", "wave-two-hands", "wave-one-hand", or "bend" which are performed by 9 different people in different scenarios with different clothing conditions as well. The video sequences are down sampled to 184*144 pixels and an average length varying from 2 to 4 seconds. This dataset contains 90 low resolution activity sequences. All the videos are having static background and running with 50 fps. We use walking, jogging and jumping sequences of Weizmann Actions dataset in this paper.

We have used templates of Mean Shift Clustering and Hu-Moments for jogging, running, walking and jumping activities as shaown in Figure 3. It is assumed that using centroid and two legs only these four activities can be identified.

B. Experimental Results

We have performed the human activity recognition experiments, with the proposed technique, on several videos, captured in outdoor and indoor environment. We have used two standard dataset namely KTH action dataset and Weizmann action dataset. In this paper, we have performed the experiments considering both indoor and outdoor scenario using KTH action dataset. But we have performed on only outdoor images of Weizmann action dataset.

1) Results on KTH dataset

Figure 4, 5 and 6 show the different frames of experimental results at different time instances on a standard KTH actions dataset. In Figure 4, first image of frame 5 shows that a human is walking. Second image of frame 5 shows the corresponding recognition result as walking with good accuracy. In Figure 5, first image of frame 10 shows that a human is jogging. Second image of frame 10 shows the corresponding recognition result as jogging. In Figure 6, first image of frame 3 shows that a human is running. Second image of frame 3 shows the corresponding recognition result as running with good accuracy.

2) Results on Weizmann dataset

To validate the robustness of our proposed method, we experimented on a standard Weizmann dataset. Figure 7, 8 and 9 shows the frame by frame result analysis of different human activity on this dataset at different time instances. In Figure 7, first image of frame 5 shows that a human is walking in outdoor environment. Second image of frame 5 shows the corresponding recognition result as walking with good accuracy.

In Figure 8, first image of frame 10 shows that a human is running in outdoor environment. Second image of frame 1 shows the corresponding recognition result as running with good accuracy. In Figure 9, first image of frame 1 shows that a human is jumping in outdoor environment. Second image of frame 1 shows the corresponding recognition result as jumping with good accuracy.

C. Result Analysis

Accuracy of proposed method is measured based on the number of frames recognized and number of frames not recognized by the following formulae:

$$Accuracy\,(\%) = \frac{No.\,of\,\,frames\,currectly\,recognized}{Total\,no.\,of\,\,video\,\,frames\,in\,a\,sequence} \times 100$$

Table 2 shows the accuracy of introduced approach over two large datasets with encouraging results; up to 95.01% of activities are recognized correctly in KTH dataset and 91.36% of activities are recognized correctly in Weizmann dataset. We have calculated the accuracy in both indoor and outdoor scenarios in the case of KTH dataset. Table 3 shows that the proposed method outperforms other existing methods.

Zhang et al. achieved 61% gait recognition accuracy over USF dataset of 4-7 activities using a simplified five-link biped locomotion human model [8]. Over indoor dataset of 5 activities, 93% accuracy is gained using the parametric model of human from image sequences using motion/texture based human detection and tracking [9]. Vega and Sarkar reported 90% accuracy using 3 actions over 71 subjects using the change in the relational statistics among the detected image features, without the need for object models, perfect segmentation, or part-level tracking [13]. Whereas, we are able

to gain upto 95% and 91% accuracy using just gait analysis over KTH and Wiezmann datasets respectively. From the experimental results it is deduced that the introduced approach is more robust and able to achieve high accuracy over large datasets by considering more activities.

(a) frame 5   (b) frame 20   (c) frame 35
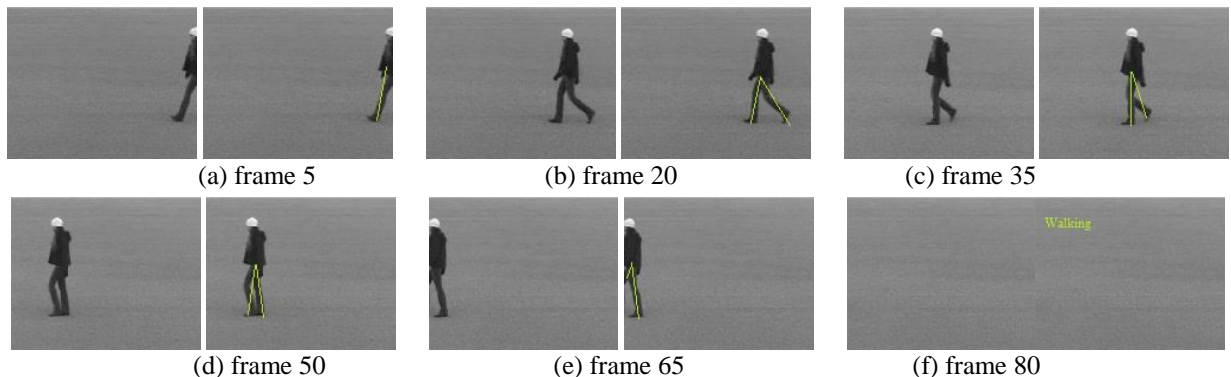
(d) frame 50   (e) frame 65   (f) frame 80

Fig. 4. Result on standard KTH dataset from of walking; first image shows input frame, second image shows corresponding output image; at the end, it recognize human activity as "Walking".

(a) frame 10   (b) frame 20   (c) frame 30

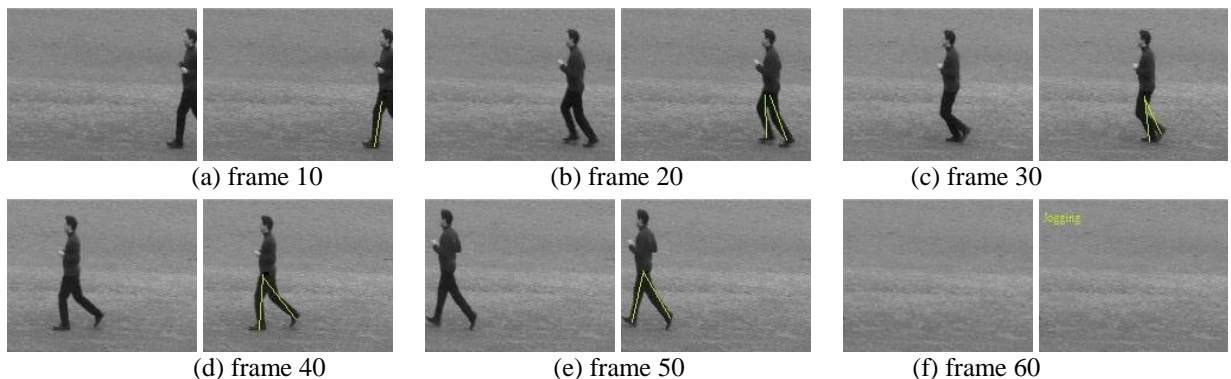(d) frame 40   (e) frame 50   (f) frame 60

Fig. 5. Experimental result on standard KTH dataset of jogging; first image shows input frame, second image shows corresponding output image; at the end, it recognize human activity as "Jogging".

TABLE 2. TABLE SHOWS THE RESULT ANALYSIS OF PROPOSED METHOD ON KTH HUMAN ACTIONS DATASET AND WEIZMANN ACTIONS DATASET ON THE BASIS OF FRAMES

| Name of Dataset	Environment condition	Human Activities	Number of Frames	Number of Frames recognized	Recognition rate
KTH Dataset	Outdoor	Walking	1443	1434	99.3%
	Indoor	Walking	1415	1383	97.7%
	Outdoor	Jogging	1525	1425	93.4%
	Indoor	Jogging	1218	1157	94.9%
	Outdoor	Running	1089	980	89.9%
	Indoor	Running	1137	1080	94.9%
				Avg. %	**95.01%**
Weizmann Dataset	Outdoor	Walking	678	650	95.8%
	Outdoor	Running	588	552	93.8%
	Outdoor	Jumping	756	642	84.5%
				Avg.%	**91.36%**
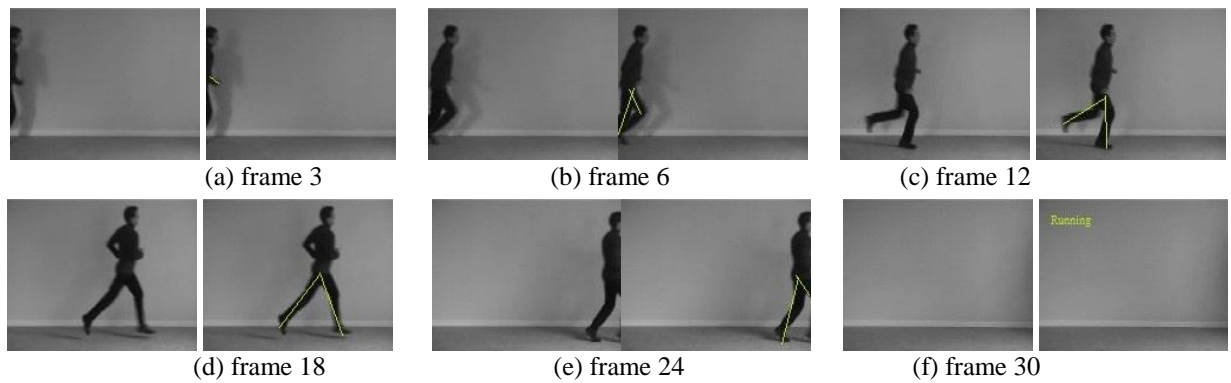
Fig. 6. Result on standard KTH dataset of running; first image shows input frame, second image shows corresponding output image; at the end, it recognize human activity as "Running".
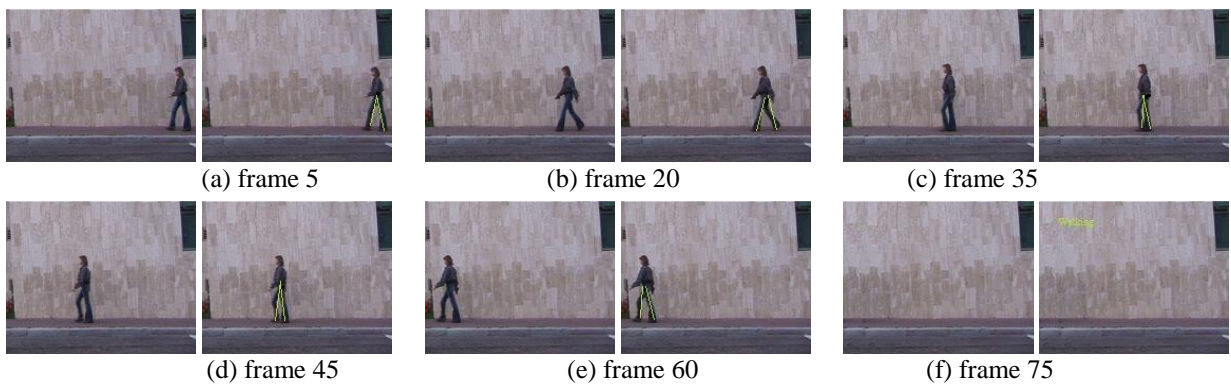

Fig. 7. Experimental result on standard Weizmann dataset of walking; first image shows input frame, second image shows corresponding output image; at the end, it recognize human activity as "Walking".
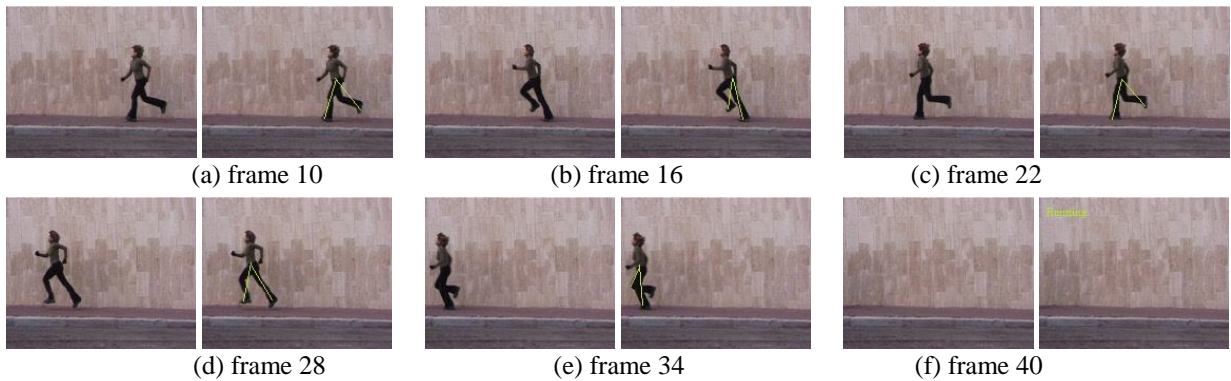

Fig. 8. Experimental result on standard Weizmann dataset of running; first image shows input frame, second image shows corresponding output image; at the end, it recognize human activity as "Running".
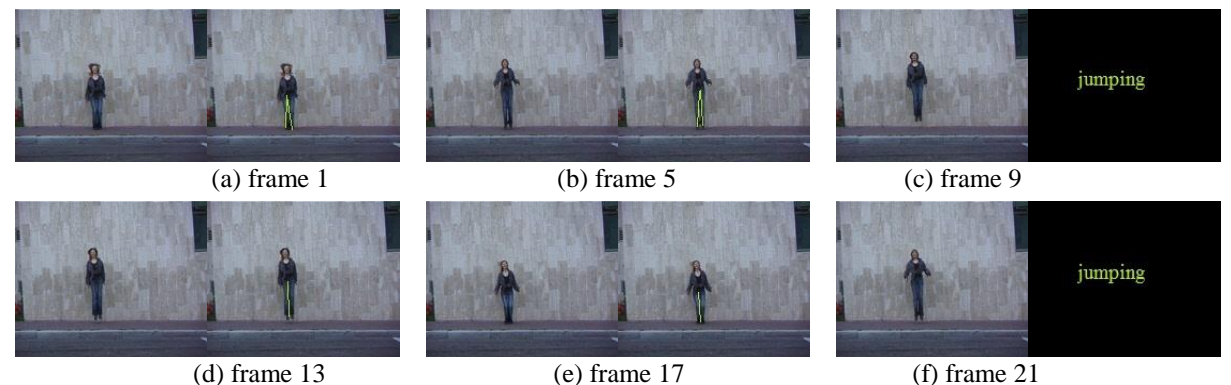

Fig. 9. Experimental result on standard Weizmann dataset of jumping; first image shows input frame, second image shows corresponding output image; at *the* end of each sub-sequence it recognize human activity as "Jumping".

TABLE 3. COMPARISON OF RESULTS WITH EXISTING METHODS

Method	Dataset	No. of Subjects	Number of Frames	Human Activities	Recognition rate
Proposed	KTH Dataset	25	7827	Walking Jogging Running	**95.01%**
	Weizmann Dataset	9	2022	Walking Jumping Running	**91.36%**
[8]	USF Dataset	75	2045	4-7 activities	**61%**
[9]	Indoor Dataset	-	9933	Standing Sitting Bending Walking Laying	**93%**
[13]	-	71	-	Walking Jogging Running	**90%**

In Figure 8, first image of frame 10 shows that a human is running in outdoor environment. Second image of frame 1 shows the corresponding recognition result as running with good accuracy. In Figure 9, first image of frame 1 shows that a human is jumping in outdoor environment. Second image of frame 1 shows the corresponding recognition result as jumping with good accuracy.

D. Result Analysis

Accuracy of proposed method is measured based on the number of frames recognized and number of frames not recognized by the following formulae:

$$Accuracy\,(\%) = \frac{No.\,of\,\,frames\,currectly\,recognized}{Total\,no.\,of\,video\,frames\,in\,a\,sequence} \times 100$$

Table 2 shows the accuracy of introduced approach over two large datasets with encouraging results; up to 95.01% of activities are recognized correctly in KTH dataset and 91.36% of activities are recognized correctly in Weizmann dataset. We have calculated the accuracy in both indoor and outdoor scenarios in the case of KTH dataset. Table 3 shows that the proposed method outperforms other existing methods.

Zhang et al. achieved 61% gait recognition accuracy over USF dataset of 4-7 activities using a simplified five-link biped locomotion human model [8]. Over indoor dataset of 5 activities, 93% accuracy is gained using the parametric model of human from image sequences using motion/texture based human detection and tracking [9]. Vega and Sarkar reported 90% accuracy using 3 actions over 71 subjects using the change in the relational statistics among the detected image

features, without the need for object models, perfect segmentation, or part-level tracking [13]. Whereas, we are able to gain upto 95% and 91% accuracy using just gait analysis over KTH and Wiezmann datasets respectively. From the experimental results it is deduced that the introduced approach is more robust and able to achieve high accuracy over large datasets by considering more activities.

V. CONCLUSIONS

An efficient human activity recognition using gait technique based on model based approach is introduced in this paper which uses Mean shift clustering algorithm and Hu-Moments to construct the activity templates. This method has a promising execution speed of 25 frames per second and good activity recognition accuracy. The experimental results demonstrate that the proposed method accurately recognizes different activities in various video frames considering both indoor and outdoor scenarios while maintaining a high recognition accuracy rate. Currently our method determines key poses of each activity independently using parametric model only. Different activity classes may give similar key poses which may cause confusion and redundancy in recognition. More discriminative key poses can be applied jointly using some more refined and sophisticated algorithms such as Support Vector Machine (SVM). We found promising recognition performance more than 95% over 3-4 activities. Experimental results suggest that the proposed method outperforms other existing methods.

REFERENCES

[1] R. Poppe, "A survey on vision-based human action recognition," Image and Vision Computing, Vol. 28, No. 6, pp. 976-990, 2010.

[2] M.K. Leung and Y.H. Yang, "First Sight: A human body outline labeling system," IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 17, No. 4, pp. 359–377, 1995.

[3] J.H. Yoo, M.S. Nixon and C.J. Harris, "Model-driven statistical analysis of human gait motion," In the Proceedings of the IEEE International Conference on Image Processing, pp. 285–288, 2002.

[4] L. Lee and W.E.L. Grimson, "Gait analysis for recognition and classification," In the Proceedings of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition, pp. 148–155, 2002.

[5] R. Tanawongsuwan and A. Bobick, "Gait recognition from time-normalized joint-angle trajectories in the walking plane," In the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 726–731, 2001.

[6] L. Wang, H. Ning, T. Tan and W. Hu, "Fusion of static and dynamic body biometrics for gait recognition," IEEE Transactions on Circuits and Systems for Video Technology, Vol. 14, No. 2, pp. 149–158, 2004.

[7] R. Zhang, C. Vogler and D. Metaxas, "Human gait recognition at sagittal plane," Image and Vision Computing, Vol. 25, No. 3, pp. 321–330, 2007.

[8] N. Nattapon, S. Nikom and K. Montri, "Model-based Human Action Recognition," In the Proceedings of SPIE, the International Society for Optical Engineering, pp. 111–118, 2008.

[9] A.F. Bobick and J.W. Davis, "The recognition of human movement using temporal templates," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 23, No. 3, pp. 257–267, 2001.

[10] J. P. Gupta, N. Singh, P. Dixit, V. B. Semwal and S. R. Dubey, "Human Activity Recognition using Gait Pattern," International Journal of Computer Vision and Image Processing, Vol. 3, No. 3, pp. 31 – 53, 2013.

[11] A.N. Rajagopalan and R. Chellappa, "Higher-order spectral analysis of human motion," In the Proceedings of the International Conference on Image Processing, pp. 230–233, 2000.

[12] I.R. Vega and S. Sarkar, "Statistical motion model based on the change of feature relationships: human gait-based recognition," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 25, No. 10, pp. 1323–1328, 2003.

[13] S. R. Dubey and A. S. Jalal, "Detection and Classification of Apple Fruit Diseases Using Complete Local Binary Patterns," In proceeding of the Third International Conference on Computer and Communication Technology, pp. 346-351, 2012.

[14] S. R. Dubey and A. S. Jalal, "Species and variety detection of fruits and vegetables from images," International Journal of Applied Pattern Recognition. Vol. 1, No. 1, pp. 108-126, 2013.

[15] W.L. Romero, R.G. Crespo and A.C. Sanz, "A prototype for linear features generalization," International Journal of Interactive Multimedia & Artificial Intelligence, Vol. 1, No. 3 2010.

[16] K.S. Kumar, V.B. Semwal and R.C. Tripathi, "Real time face recognition using adaboost improved fast PCA algorithm," arXiv preprint arXiv:1108.1353, 2011.

[17] S. R. Dubey and A. S. Jalal, "Adapted Approach for Fruit Disease Identification using Images," International Journal of Computer Vision and Image Processing, Vol. 2, No. 3, pp. 44-58, 2012.

[18] H. Bolivar, A. Pacheco and R.G. Crespo, "Semantics of immersive web through its architectural structure and graphic primitives," International Journal of Interactive Multimedia & Artificial Intelligence, Vol. 1, No. 3, 2010.

[19] S. R. Dubey, P. Dixit, N. Singh, and J.P. Gupta, "Infected fruit part detection using K-means clustering segmentation technique," International Journal of Artificial Intelligence and Interactive Multimedia. Vol. 2, No. 2, pp. 65-72, 2013.

[20] N. Singh, S. R. Dubey, P. Dixit and J.P. Gupta, "Semantic Image Retrieval by Combining Color, Texture and Shape Features," In the Proceedings of the International Conference on Computing Sciences, pp. 116-120, 2012.

[21] S.J.B. Castro, R.G. Crespo and V.H.M. García, "Patterns of Software Development Process," International Journal of Interactive Multimedia & Artificial Intelligence, Vol. 1, No. 4, 2011.

[22] S. R. Dubey and A. S. Jalal, "Robust Approach for Fruit and Vegetable Classification," Procedia Engineering, Vol. 38, pp. 3449-3453, 2012.

[23] R.G. Crespo, S.R. Aguilar, R.F. Escobar and N. Torres, "Dynamic, ecological, accessible and 3D Virtual Worlds-based Libraries using OpenSim and Sloodle along with mobile location and NFC for checking in," International Journal of Interactive Multimedia & Artificial Intelligence, Vol. 1, No. 7, 2012.

[24] V.B. Semwal, V.B. Semwal, M. Sati and S. Verma, "Accurate location estimation of moving object in Wireless Sensor network," International Journal of Interactive Multimedia and Artificial Intelligence, Vol. 1, No. 4 , pp. 71-75, 2011.

[25] C.J. Broncano, C. Pinilla, R.G. Crespo and A. Castillo, "Relative Radiometric Normalization of Multitemporal images," International Journal of Interactive Multimedia and Artificial Intelligence, Vol. 1, No. 3, 2010.

[26] P. Siirtola and J. Röning, "Recognizing Human Activities User-independently on Smartphones Based on Accelerometer Data," International Journal of Interactive Multimedia & Artificial Intelligence, Vol. 1, No. 5, pp. 38-45, 2012.

[27] Y. Cheng, "Mean shift, mode seeking, and clustering," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 17, No. 8, pp. 790-799, 1995.

[28] M.K. Hu, "Visual pattern recognition by moment invariants," IRE Transactions on Information Theory, Vol. 8, No. 2, pp. 179-187, 1962.

[29] M.G. Gazendam and A.L. Hof, "Averaged EMG profiles in jogging and running at different speeds," Gait Postures, Vol. 25, No. 4, pp. 604–614, 2007.

[30] S. R. Dubey, N. Singh, J.P. Gupta and P. Dixit, "Defect Segmentation of Fruits using K-means Clustering Technique," In the Proceedings of the Third International Conference on Technical and Managerial Innovation in Computing and Communications in Industry and Academia. 2012.

[31] N. Singh, S. R. Dubey, P. Dixit and J.P. Gupta, "Semantic Image Retrieval Using Multiple," In proceeding of the 3rd International Conference on Technical and Managerial Innovation in Computing and Communications in Industry and Academia. 2012.

[32] K.S. Kumar, V.B. Semwal, S. Prasad and R.C. Tripathi, "Generating 3D Model Using 2D Images of an Object," International Journal of Engineering Science, 2011.

[33] V.B. Semwal, V.B. Semwal, M. Sati and S. Verma, "Accurate location estimation of moving object in Wireless Sensor network," International Journal of Interactive Multimedia and Artificial Intelligence, Vol. 1, No. 4 , pp. 71-75, 2011.

[34] K.S. Kumar, S. Prasad, S. Banwral and V.B. Semwal, "Sports Video Summarization using Priority Curve Algorithm," International Journal on Computer Science & Engineering, 2010.

[35] S.R. Dubey, "Automatic Recognition of Fruits and Vegetables and Detection of Fruit Diseases", M.Tech Thesis, GLAU Mathura, India, 2012.

[36] M. Sati, V. Vikash, V. Bijalwan, P. Kumari, M. Raj, M. Balodhi, P. Gairola and V.B. Semwal, "A fault-tolerant mobile computing model based on scalable replica", International Journal of Interactive Multimedia and Artificial Intelligence, Vol. 2, 2014.

[37] S.R. Dubey and A.S. Jalal, "Automatic Fruit Disease Classification using Images", Computer Vision and Image Processing in Intelligent Systems and Multimedia Technologies, 2014.

[38] D.D. Agrawal, S.R. Dubey and A.S. Jalal, "Emotion Recognition from Facial Expressions based on Multi-level Classification", International Journal of Computational Vision and Robotics, 2014.

[39] S.R. Dubey and A.S. Jalal, "Fruit Disease Recognition using Improved Sum and Difference Histogram from Images", International Journal of Applied Pattern Recognition, 2014.

Jay Prakash Gupta received his MTech in CSE in 2012 from GLA University Mathura, India and BTech in CSE in 2010 from B.S.A. College of Engineering and Technology Mathura, India. He is working as a System Engineer in Infosys Limited, Pune, India. His research interests include image processing, computer vision and pattern recognition.

Pushkar Dixit is currently working as an Assistant Professor in Computer Science and Engineering, FET Agra College, Agra, India. He received his MTech in Computer Science from GLA University Mathura, India and BTech in Computer Science from FET Agra College Agra, India. He has 3 year of teaching and research. His research interests include image processing, network security and computer networks.

Nishant Singh received his MTech in CSE in 2012 from GLA University Mathura, India and BTech in CSE in 2010 from B.S.A. College of Engineering and Technology Mathura, India. He has 1 year of teaching and research experience and currently, he is working as an Assistant Professor in Department of Computer Engineering and Applications, Poornima Institute of Engineering and Technology, Jaipur, India. His research interests include image processing, computer vision and pattern recognition.

Vijay Bhaskar Semwal has been severed as a PCC member in several International conferences and Journals. He was also head for 5th IEEE conference on big data at COER Roorkee. He is carrying more than 5 year of academic and industry experience. He served for various top organizations like Siemens, Newgen etc. Earlier he received his M.tech in wireless sensor domain and was awarded by gold medal. His major research work interest in Wireless Sensor Network, Artificial Intelligence, Image Processing, Computer Network &Security, and Design & Analysis of Algorithm, Machine Learning, and Information Retrieval Soft Computing