International Journal of Research in Advent Technology, Vol.7, No.4, April 2019
E-ISSN: 2321-9637
Available online at www.ijrat.org

Human activity recognition using foots movement patterns

H. Chidananda[1], Dr.T Hanumantha Reddy[2]
Department[1,2] Department of Computer Science &Engineering,
Affiliation name[1,2] Rao Bahadur Y Mahabaleswarappa Engineering College,
Email: chidably999@gmail.com[1], thrbly@yahoo.com[2]

Abstract- This work proposes a real time rule based human activity recognition system using foots movement patterns in a video sequence. The activities like walking, running and jogging are evaluated using the angle and the distance between the extreme points of right and left foots. These extreme points of the lower part of human model (rear end point of back foot and front end point of front foot) are used as main feature points. The aim of this work is to automate the system that understands the human activities in a video sequence just enough with one complete cycle of actions of the activity. A rule based classifier is used to classify the activities in a video sequence. The proposed work has been implemented on standard KTH datasets and achieved an overall 95.10% of efficiency without using any complicated time consuming classification algorithms.

Keywords - Activity recognition, foots patterns, walking, jogging, running, foot movements, human activities.

1. INTRODUCTION

As of today, vision based human activity analysis and understanding has been a challenging and most important aspect in many applications like elderly health care systems, surveillance systems, intelligent transport systems, human-machine-interaction system and many more. There are many approaches exists but all these belongs to either of the two most important categories like model-depended and model free approaches. A survey [1] has been made on vision-based human action recognition by poppe that one can easily watch the walking style and its patterns, people can also discover the whole body style in gait but they can even understand the motion of the particular segments of the human body. The first category i.e. gait understanding based on model approach consider motion of the human structural design and using these motion a gait patterns will be constructed using the design parameters. Whereas the other approaches do not try to get back the design patterns of human movements. The main parameters adopted to represent gait are shape movements, width, height and templates of images.

Leung & Yang [2] has addressed the basic issues of segmentation, human tracking and classification of human body segments from a human contour. Generally human body design consists of five U-pattern strips, a trunk of the human body, different joints and mid-points. Along with the general two dimensional structure, a view based knowledge has been described for various basic postures like side-view kneeling, side horse moments which helps in understanding the method used. The human contour segmentation is formed by discovering moving segments. Yoo et al., [3] has evaluated the angle between knee and hip joints from the human body silhouette by examining linear regression. To these series of angles a trigonometric polynomial interpolation exercises will be equipped and the features acquired will be helpful for activity understanding.

The human contour is separated into segments which equivalent to various body segments, and an every segment fit elliptical shapes to obtain the structure of human model is described in[4]. Accurate reconstruction of human structure is very much confide on extraction of human contour. But the evaluated features need not be trustworthy.

Bobick and Tanawongsuwan [5] remodeled the human model by perusing the three dimensional sensors deployed on joints. As their method recognized walking action only, so human interaction is very much required. But opposed to this, the proposed method understands three human activities and achieved good results. A two dimensional of human cone was built by Wang et al... They have tracked the person who walks under the abstract framework and modeled static and non-static parameters from various human body segments for understanding gait activity [6]. Their strategy was to combine both static and non-static parameters to get good gait activity recognition. However using both static and non-static parameters requires more calculations and processing time.

A simple 5-link-biped action design for understanding gait activity has been proposed Zhang et al. [7]. They captured important gait parameters from series of frames and these are then deployed to train up HMM for understanding of the activity. A vision based human activity understanding method was proposed by [8] adopting the human parametric-model from series of frames using its texture/movements. They deployed the texture/moments of whole human body segments but the proposed work uses the foot patterns of the human body which is good in time consumption.

Davis & Bobick used Motion-Enegy-Images (MEI) and Motion-History-Images (MHI) to understand human moments in a series of frames [9]. The moving frames of images in a series were determined by performing difference between consecutive images in a series of frames and used some threshold level to convert into a twofold value. The MEI is formed by collecting these action frames within time. After then these MEI are enriched to MHI wherever the value of the pixel is corresponding to the time of action at the point. The motion-based parameters are retrieved out of MHI & MEI and then these are adopted for understanding the activity by utilizing layout coordinating. Since this strategy depends on whole layout coordinating rather than the main walking example of the legs, it will not take the benefit of latest development but the proposed work implemented the corresponding using the walking process. Latest walking investigation reports or action understanding advices the walk pattern is an exclusive individual typical feature with rhythm and cyclical as in [10]. Challappa and Rajagopalan [11] portrayed a superior-order phantom investigation method for distinguishing persons by perceiving human action like strolling/running. In this strategy, walk dimensions were resolved in each frame.

Vega and Sarkar [12] offered a novel portrayal plot for vision based movement investigation utilizing only the changes in the relational insights among the distinguished features of the image, without the requirement for object models, consummate division, or tracking part-level. They displayed the relational insights utilizing the likelihood that an irregular gathering of highlights in a picture would show a specific connection. To decrease the authentic combinatory of these relational disseminations, they performed them in a Space of Probability Functions (SoPF). Distinctive movement sorts clear out various follows in this space. They additionally showed and assessed the viability of that portrayal with regards to perceiving people from gait. In any case, there strategy requires numerous cameras from various perspectives to demonstrate multi-view acknowledgment framework which requires additional setup and furthermore calculation, while the proposed approach can accomplish higher recognition rate from single 2d camera using the angle and the distance between the extreme points of the right and left foots in only one complete cycle of the foot movements. A few different methodologies and highlights utilized as a part of [13-25] might be fixing with walk investigation to estimate the human activities. Human action acknowledgment utilizing cell phones is likewise contemplated [26] however its acknowledgment rate can be enhanced utilizing walk examination with additional time proficiently.

2. PROPOSED METHODOLOGY

The proposed technique of human activity recognition is based on the foreground extraction of frames, tracking of the human, extraction of feature points and recognition of the activity. The proposed system's framework is shown in figure 1, uses foot movements as a pattern points to distinguish three essential human activities of walking, jogging, and running. The proposed system accepts the video as input from the human action database and frames are extracted for one cycle of the activity from the video sample. The two extreme points on the two foots (rear end point of back foot and front end point of front foot) are used as feature points. The angle and distance between these points are evaluated for one complete cycle of foot movements and classified using rule based classifier to recognize the activity. A rule based classifier adopts three threshold levels to classify the activities like walking, running, and jogging. The performance of the proposed method has been evaluated experimentally on MATLAB using the standard KTH action datasets.

2.1. Foreground Extraction

The proposed system uses a video sequence of human activity as an input. The video contains various continuous frames out of which only frames that are part of one complete cycle of the activity are extracted for processing and evaluation. For each frame a foreground extraction [27] is applied for further processing. In any case, these frames contain a few noise elements which may prompt incorrect foreground area subtraction. So unwanted noise parts are removed. Some of the small noise content are removed by using morphological image processing tools such as Erosion, Dilation, or Gaussian Filters.

2.2. Feature Extraction

This work employed a two important extreme points, one on each foot on the lower part of the human model as feature points as shown in figure 2. These points are determined using the algorithm1. These points are used to track the foot movement patterns as the two legs continuously move in one distinct pattern as in the case of walking, running and jogging activities.

2.3. Activity Recognition

Steps for recognition of human activities using rule based classifier are given in algorithm2. Evaluation process of the extracted frames to recognize the human activity in a video sequence is as follows:

Find the Euclidean distance and angle in radians between the two extreme points one on each foot (front and back) fp1 at (x_1, y_1) and fp2 at (x_2, y_2) as given in equation (1) and (3).

$$v= \sqrt{((y_2 - y_1)^2 + (x_2 - x_1)^2)} \qquad (1)$$

Where v is the distance between the two points.

$$||P|| = [v_1, v_2, v_3 \ldots v_n] \qquad (2)$$

Where ||P|| is the feature distance vector formed on each frame for one cycle of the activity.

$$r = \operatorname{atan}\left(\frac{y2-y1}{x2-x1}\right) \qquad (3)$$

Where r is the angle in radians between the two points fp1 and fp2 with respect to the horizontal axis.

$$||R|| = [r_1, r_2, r_3 \ldots r_n] \qquad (4)$$

Where ||R|| is the feature angle vector formed on each frame for one cycle of the activity.

$$Y = \frac{1}{n}\sum_{i=1}^{n}(Pi * Ri) \qquad (5)$$

Where Y is a classifier value, based on this the activity is recognized.

Activities are recognized using the rule based classifier as given in table 1.

2.4. Threshold levels used for classification

This work uses different threshold levels for recognition of different activities.

Walking Activity: In this activity, the distance and angles in each frame changes as the two legs moves. In walking case, there is a chance of getting more number of zero radians as the two legs reaches on the ground so the classifier value Y is less when compared to the other activities like jogging and running. So, the threshold value used for this activity is in the range 1 to 3. This range falls under threshold level 1.

Jogging Activity: In jogging activity, the distance vector is less as compared in running activity. So the threshold is set for jogging activity in between 4 to 8. This falls under threshold level 2.

Running Activity: In this the distance vector is more as compared to both walking and jogging activities. So the threshold level is set for the running activity is 9 to 12, which is in threshold level 3. The figure 3 shows the graph for the three activities.

Algorithm 1: Foot pattern points extraction. Input: Video Sequence V Output: Foot pattern points fp1 and fp2	
1	Input the sample activity video V
2	f_n ← extract n frames from v for one cycle of the activity.
3	for each n frames, traverse and process: fb_i ← binarize(fbb_i), fbb_i ← bob(fb_i) ← foreground extraction [27] ← human tracking [28]
4	c ← contour(fbb_i)
5	Find h and w to locate scanning area. h ← height(fbb_i) h ← h/4 (from the ground to focus only on the lower part of the human body) w ← width of (fbb_i)
6	Scanning area S ← h * w
7	fp1 ← scan vertically towards right for first contour pixel from the bottom left of the scan area S to get fp_1, repeat this towards right $\|\|^{ly}$ fp2 ← scan vertically towards left for first contour pixel from the bottom right of the scan area S to get fp_2
8	End

Algorithm 2: Activity Recognition Input: fp1 and fp2 Output: Activity Recognition	
1	||P|| ← distance vector between fp1 and fp2 as in eq (1) and (2)
2	|R||← angle vector (in radians) between fp1 and fp2 as in eq (3) and (4)
3	Y ← Classifier value as in eq (5)
4	if Y falls in threshold level 1 (as given in Table 1), then the activity is Walking, if Y falls in threshold level 2, then the activity is Jogging and if Y is in threshold level 3 then the activity is Running.
5	End

3. RESULTS AND DISCUSSIONS

This section analyses the various aspects of the proposed method. The foot movement patterns are formed by using the algorithm 1. The evaluation of the foot movement patterns are processed and analyzed for better understanding of the activity by using the algorithm 2.

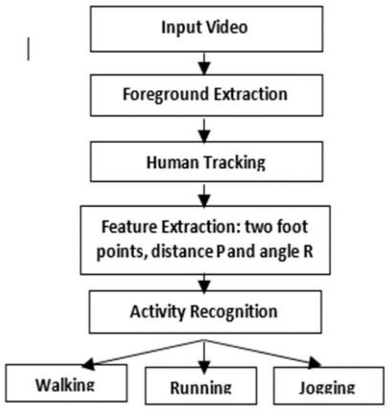
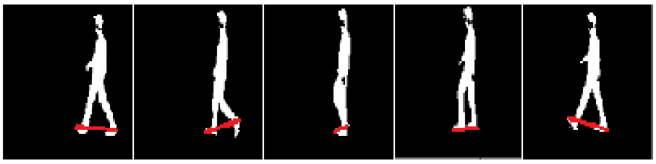


Figure 1: Framework of Human Activity Recognition

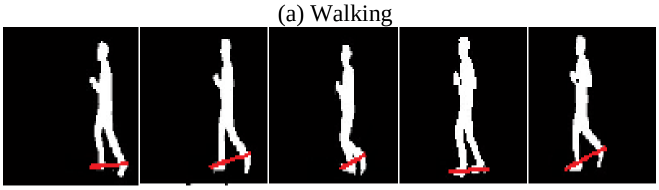
3.1. Data Set Used

In order to evaluate our proposed approach of human activity recognition, the work is implemented on MATLAB using standard KTH Human Actions dataset (http://www.nada.kth.se/cvap/actions).

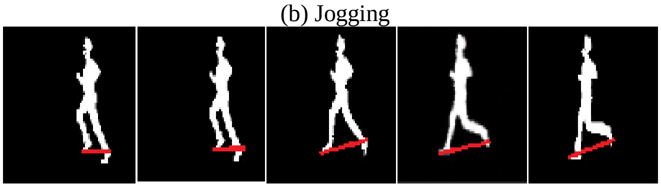
KTH Human Actions dataset: KTH video dataset utilizes three sorts of human activities, for example, Walking, Jogging and Running which were performed by 20 subjects in various conditions with various dress code.
The visual sequences are down sampled to 160*120 pixels and an average video length varying from 4 to 41 seconds. This dataset contains 2391 activity sequences. All videos are having static background with 25 fps. The proposed work uses two leg moments for recognizing jogging, running and walking activities as shown in figure 3.



(a) Walking



(b) Jogging

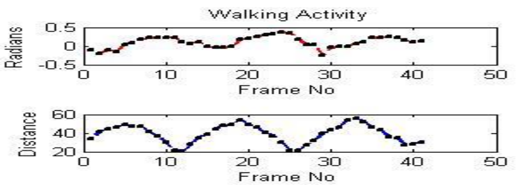


(c) Running

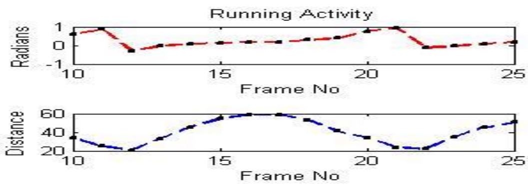
Figure 2. Silhouette of foot movement patterns for (a) Walking, (b) Jogging and (c) Running

Table 1 Threshold levels used in Rule Based Classifier.

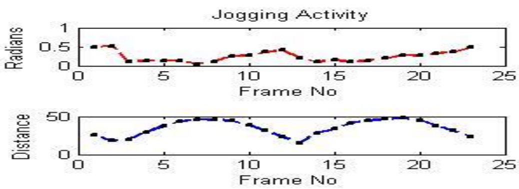
		Threshold Level(Th)	Rule	Activity
		Th1	1 to 3	Walking
Y	{	Th2	4 to 8	Jogging
		Th3	9 to 12	Running



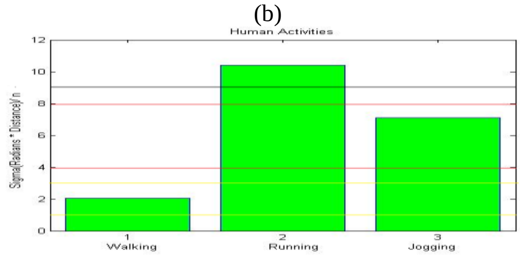
(a)



(b)



(c)



(d)

Figure 3: Graphs representing the three activities for (a) Walking, (b) Jogging (c) Running and (d) the threshold levels used to classify the activities.

Table 2. Evaluation of our method on KTH Human Actions dataset

Dataset Used	**Human Activity**	**No. of Frames Used**	**No. of video samples**	**% rate**
KTH	Walking	35 to 42	20	98.5 %
	Jogging	23 to 28	15	94.5 %
	Running	26 to 30	15	92.3 %
AVG %				95.10 %

3.2. Experimental Results

The proposed work have conducted the individual activity classification experiments using several videos and standard dataset namely KTH human actions dataset.

3.3. Results on KTH dataset

The proposed work has been tested on KTH dataset and the results are shown in table 2. The frames of these video dataset are extracted for one view i.e. first frame to the last frame in the view of the three activities. Euclidean separation and angle in radians of the two (right and left) foots points are used to determine the activity as shown in figure 2. The final graph for the three activities and its threshold levels are shown in figure 3. This method has been compared with other works [8], [9] and [10] is shown in table 3.

Table 3: Overall Recognition Rate

Method	**Dataset**	**No. of Subjects**	**No. of frames used**	**Human Activity**	**% Rate**
Our	KTH	30	23 to 42	Walking, Jogging and Running	95.10%
[8]	KTH	25	7827	Walking, Jogging and Running	95.01%
[9]	USF Dataset	75	2045	4-7 activities	61%
[10]	Indoor Dataset	-	9933	Walking, Jogging and Running	90%

CONCLUSION

The proposed work has evaluated the foot movement patterns of walking, jogging and running activities of The proposed work has evaluated the foot movement patterns of walking, jogging and running activities of video sequences to understand and recognize the vision based human activities. In each of these activities the length and the angle vector values of foots movement pattern for only one complete cycle of the activity are considered and classified the activities accurately using rule based classifier. The work has been tested on MATLAB using a challenging KTH human actions datasets and achieved overall classification rates up to 95.10% without using any complicated and time consuming classification algorithms.

Future research work shall focus on the classification of speed, slow and fast walking/jogging/running activities on different terrains.

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 Multi temporal 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]. Z. Tang, Z. Miao, Y. Wan and J. Li, "Automatic foreground extraction for images and videos," 2010 IEEE International Conference on Image Processing, Hong Kong, 2010, pp. 2993-2996.

[28]. T. Zhang, "Video security with human identification and tracking," 2014 IEEE International Conference on Multimedia and Expo Workshops (ICMEW), Chengdu, 2014, pp. 1-6.