*International Journal of Application or Innovation in Engineering & Management (IJAIEM)*
**Web Site: www.ijaiem.org Email: editor@ijaiem.org**
**Volume 7, Issue 8, August 2018 ISSN 2319 - 4847**

# Activity Recognition Using Particle Network

**Ali Parsa Sirat [1]**
[1] *Master Student,* Sharif University of Technology, Iran

**ABSTRACT**
***Nowadays, human activity recognition is considered as an important field in machine vision. It has a lot of applications in video surveillance, person recognition and prediction. Human activity recognition is a challenging task because of its complexity and highly diversity of actions. Most of the previous methods used tracking algorithms to track the person and then recognize him. In this paper we propose a new method which can be used without tracking the person. In this method, we consider particles on frame, and then by using optical flow algorithm on these particles we obtain some specific features. Using this feature vector can help us to recognize the activity of people. Finally, we will have a feature vector for each frame of activity and by using classification methods such as Adaboost we are classifying the activities. We evaluate this method on Weizmann dataset. The proposed method in addition of having high accuracy, is independent of the need for video tracking algorithms and their constraints***
**Keywords:** Human activity recognition, Particle filter, Particle network, Mixture of Gaussian, AdaBoost.

## 1. INTRODUCTION

Human activity recognition is an important subject in machine vision field. In recent years activity recognition, due to its' vast application in security, surveillance and entertainment attracts the attention of a lot of researchers. Video surveillance can be used in a variety of environments: banks, metros, airlines for people counting, human activity prediction. Up until now various methods proposed for categorizing the methods for activity recognition. Both proposed categorizing method in [1], [2] is based on features. In [1] they divide the process to three sections: 1- representation, 2- segmentation, 3- recognition, in other hand [2] divide the process in 2 sections: 1- representation, 2- the subsequent classification process. Also, in terms of pattern recognition models, some papers have done valuable work in this area [3], [4]. Overall in the activity recognition image representation partition in two categories: global representations and local representations. Overall in the activity recognition image representation partition in two categories: global representations and local representations.

In global representation all the information from sequence of images in video is considered. Since the global representation consider all the information, it is a strong method for recognition however this method is sensitive to occlusion and different viewpoints. In [3] an unsupervised method [6][7] for human activity recognition specially for Sports videos is proposed. At the beginning four major points: 1- head center, mass center, left end of leg and right end of leg, then a statistical analysis for these four points using human silhouette, is performed. They used a camera estimation and object localization method.

In [23][9], energy-base approach is introduced for detecting abnormal behavior. They used adaptive optical flow model to calculate the energy of each frame. They extract foreground then due to the presents of noise on videos they used Horn-Shunk algorithm to extract the features from moving objects on the image [10].

In contrast in local representation the local information is extracted separately. In this method, the major points will be detected then a set of neighborhood pixels will be considered [11][12] [13].

Combination on local regions provides the final representation of image. This method in comparison to the global representation has less sensitivity to the noise and occlusion. This method is independent of foreground extraction and tracking.

In following we discuss some methods which are based on local representation. In [14][15] video representation construct by Haris feature extraction and integrate with SVM classification for recognition. In this method, information is extracted only from the important part of the images and it does not consider the relation between regions.

Authors in [16] represents human action by pose without considering complexity of dynamics representation. A pose descriptor called Histogram of Oriented Rectangles introduced for representing and recognizing human actions in videos. They also incorporate velocity descriptor for some cases which HOR is not sufficient by itself to differentiate activities such as "jogging" and "running".

Patterns of motion also is presented in [17] for action recognition. They used mid-level motion features which are extracted using low-level optical flow information. This method is based on local regions of the image regions.

Different optimization method is used in [18],[19].

An adaptive algorithm is also proposed in [20], [21] Modeling human behavior accurately, improves system planning in the different area such as advertising, marketing and even power system where demand response algorithm highly affected by uncertain human usage patterns[22][23].

Activity recognition is also an important field in robotics such as service robots in which robot by recognizing activity of elderly or deaf people can service them. By appearance of Kinect some 3D approaches are also used for human recognition. Weight adaption method for increasing recognition is also introduced for increasing accuracy.

Some methods used machine learning for categorizing [24]. They introduced a deep learning method for which in addition of categorizing, they understand the image.

Reza et al proposed mathematical method for analyzing vehicular model applied in transportation system which could be effective in future system modelling [25][26][27][28]. Some energy managements methods are also introduced in [29]. A novel computational method was presented in [30][31] to enhance recognition method using analytic software.

In this paper we introduced a new method for activity recognition in videos. In our proposed method, we separate k consecutive frames as a segment then by using Gaussian filter [32] we smooth each frame in each segment. In next step we used MOG to extract a person silhouette. This silhouette helps us to find velocity, accelerate, angle, and angular speed characteristic of a person. Then we create a person model using the above-mentioned representors. For classifying different activities based on these characteristics we used Adaboost classifier, which is one the well introduced classifier in this field.

The reminder of this paper organized as follows. In Section II we explain an overview of proposed activity recognition algorithm with each of its sub algorithms. Experimental setup and results are presented in Section III. Finally, conclusion is drawn, and future work proposed in Section V.

## 2. PROPOSED METHOD

In this section we describe our proposed method for human activity recognition. Our method is described in three subsections 1- preprocessing, 2- feature extraction 3- activity recognition. In preprocessing section

We perform preprocessing to improve image data in a way that removing image distortion which may cause errors in recognition. In second step instead of using all image information which most of them may be irrelevant we just extract features. In the last step these features help us to distinct activities.

### 2.1 Preprocessing

One of the key step in this domain is image preprocessing. We exploit Gaussian filter on image to decrease the noise on image. As we mentioned previously we do not use tracking and detection for extracting moving object since it may cause some errors by itself before the recognition part.

We put particles all over the image after that we extract foreground using MOG (Mixture of Gaussian) method. Thus, just by updating particles located on foreground we can extract the features.

The overall pseudocode of the algorithm is shown in Table 1.

### 2.2 Feature Extraction

Selecting suitable features is a crucial part of each recognition algorithm due to its important role to distinct the different classes. If features have a correlation it may cause redundant information which increase the time complexity of algorithm. Features should be meaningful and each of them by itself should be able to distinct at least two classes. Thus, the features should be discriminative with respect to the posture and motion of human body. Since we locate particles in each frame, it is needed to extract from the movement of particles on sequence frames. Thus, we separate the frames in video to segments, then for each segment we extract a feature vector. Based on features vectors belonging to the segments we recognize different activities.

**Table -1**: Features of particles on image

| |
|---|
| **Procedure PNAR**(***seqFramVideo***) |
| Input**-** s*eqFramVideo: sequence of frames from video* |
| Output**-** *class: class of activity which occurs in this video* |

```
S_{1..i} = Segmentation(frames);
- Preprocessing
filteredFrames = GaussianFilter(frames);
foregrounds = MOG(frames);
- Feature Extraction
Locate particles in first frame of each segment;
Applying HS on frames;
FeatureVector = calculating feaures: X, Y, ΔX, ΔY, V,
Angle, ΔAngle;
UpdatedFeatureVector = Update (FeatureVector,
frames);
-Recognition
AdaboostTrainModel = Train(Adataboost,
trainSetFeatureVectors);
ClassLabel = AdaboostTestModel (testFeatureVector);
return ClassLabe.l
```

To obtain these feature vectors we allocate some articles all over the frame. The number of particles is considered as N_(w,h) in which w is the width of the frame and h is the height of the frame. For each 10×10 window of frame we put a particle at the middle of window. In other words, for each 10×10 pixels we assign a particle to represent that window.

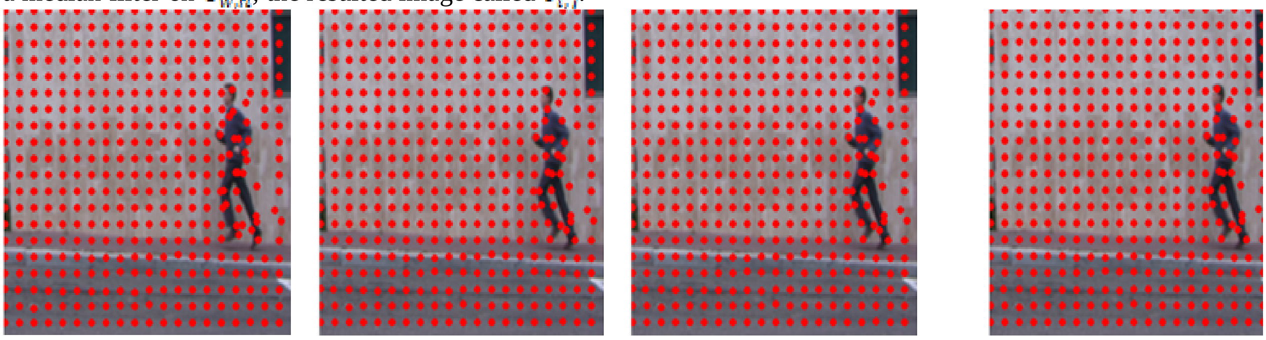
**Table -2:** Features of particles on image

| Symbol | Description |
|---|---|
| X | Position of particle in X axis |
| Y | Position of particle in Y axis |
| ΔX | Change of position in X direction |
| ΔY | Change of position in Y direction |
| V | Velocity |
| Angle | Angel relative to X axis |
| ΔAngle | Change of angle |

This particle shows the behavior of its corresponding pixels. For each particle we consider position and speed and angle of particles in each frame. In Table 2 we show these features.
Considering this feature, we can create a vector of features for all particles in frame thus we will have at the end a matrix of features to help us recognition.

### 2.1 Motion filtering

In previous section we describe what is the particles features and how we should calculate them based on foreground.
Now we should calculate the characteristic of particles on foreground of the frames. To do this purpose we applied Horn and Schunck (HS) optical flow algorithm. This algorithm shows how the flow is expected to vary across the image thus it gives us the velocity image which we called it $O_{w,h}$. This image contains velocity in direction of X and Y. As paper **Error! Reference source not found.** noted applying median filter is a key to performance gains. We applied a median filter on $O_{w,h}$, the resulted image called $\vartheta_{i,j}$.



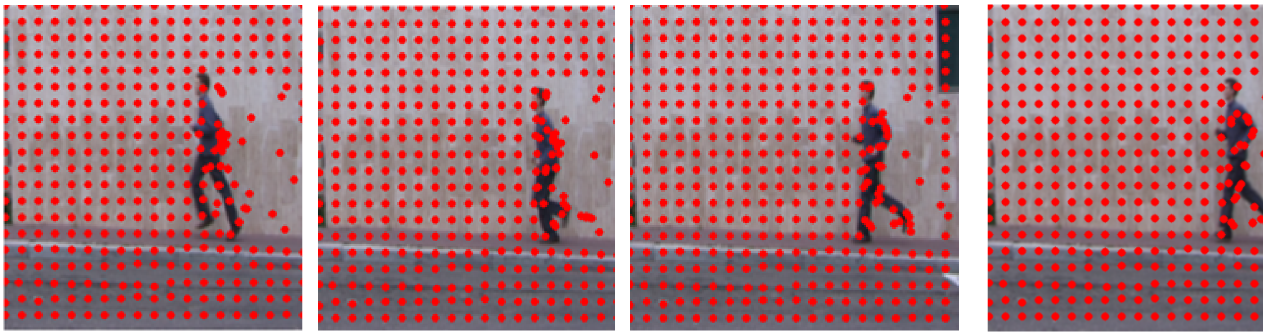


**Fig -1**: Particles are following person movement in dataset Weizmann [25]

### 2.3 Updating features of particles

In this part we describe how we update the features of each particle. The speed of each particle is the average of speed of its corresponding pixels which obtained by HS. In addition, change in position $\Delta X_k$ of $k^{th}$ particle considered as an average of $\vartheta^x_{i,j}$ which is the corresponding pixels to $k^{th}$ particle:

$$\Delta X_k = \sum_{i\epsilon k}\sum_{j\epsilon k} \vartheta^x_{i,j}$$

The similar method is used for obtaining $\Delta Y_k$. Thus, we have:

$$\Delta Y_k = \sum_{i\epsilon k}\sum_{j\epsilon k} \vartheta^y_{i,j}$$

The current position of each particle $X_k$ and $Y_k$ can also be obtained by the sum of previous position of particle in addition to the change of that:

$$X_k curr = X_k prev + \Delta X_k curr$$

Similarly, we can calculate the current position of particle in Y axis.

$$Y_k curr = Y_k prev + \Delta Y_k curr$$

Exploiting the previous calculations, particles will follow the movement of person in the image. In following we provide some sample of this particle movement:

The value of particle's angle is measured using velocity of in direction of X related to velocity in Y direction. The angel of velocity of each particle in $t^{th}$ frame is:

$$Angle(t)_{i,j} = \frac{\vartheta^x_{i,j}}{\vartheta^y_{i,j}}$$

In which t is the number of frame i,j is the pixel position in image.

$$Angle(t)_k = \sum_{i\epsilon k}\sum_{j\epsilon k} Angle(t)_{i,j}$$

To get the change of angle for each particle we used following formula:

$$\Delta Angle(t)_k = Angle(t-1)_k - Ang\,le(t)_k$$

These characteristics are calculated for each particle in each frame.

This way the particles follow the movement in the frames. Figure 1 shows some sample of particles which are following person movement in Weizmann dataset [25]. For each frame we calculate the features the concatenation of features in a sequence of frame produces feature vector. These vectors in next steps will be used for training and testing the activities.

### 2.4 Updating features of particles

There exists so many classifiers for recognition however Adaboost which is appendix for adaptive boosting in recognition showed a great achievement over classification. In this algorithm the output of some weak classifier with a weighted sum method combine with each other and result a strong recognition. Since it is the combination of weak classifier time complexity of that is low. We should provide a feature vector for training this classifier, thus we provide the previous mentioned feature vector in section 2.3 for training. In this case we have a consequence of segments in each video, we extract feature vector for each of them, then classifier will be trained by a set of feature vector of a video. Finally, in test video by its feature vector and detect the label of the activity it belongs to.

## 3. EVALUATION

We have tested our algorithm on Weizmann human action dataset [3].

This dataset contains 81 videos from 9 persons and each of them perform 10 different actions including: “walk”, “run”, “jump”, “gallop sideways”, “bend”, “one-hand wave”, “two-hands wave”, “jump in place”, “jumping jack”, “skip”. The resolution of frame in this dataset is 180×144. Among these activities we test our result on three of them “bend”, “walk” and “run”. We perform “leave one out” cross validation to evaluate the accuracy of our algorithm. Since this dataset contains single-action video sequence we report per-video classification, thus the label of video is the majority voting to the sequence of segments.

**Table -3**: The confusion matrix between run and jump

| | Running | Jumping |
|---|---|---|
| Running | 0.75 | 0.25 |
| Jumping | 0.19 | 0.81 |

**Table -4**: The confusion matrix between run and bend

| | Running | Bending |
|---|---|---|
| Running | 1.0 | 0.0 |
| Bending | 0.0 | 1.0 |

**Table -5**: The confusion matrix between jump and bend

| | Jumping | Bending |
|---|---|---|
| Jumping | 0.88 | 0.12 |
| Bending | 0.0 | 1.0 |

We provide the confusion matrix for pair of activities, as you can see we get the accuracy high 100% for discriminating run and bend, for run and jump we reach the accuracy of 75% and finally for jump and bend we get 88%. The accuracy of proposed method over all activities is also 83.3%.

## 4. CONCLUSION

In this paper, we introduced a new method for human action recognition using particle network. We locate particles all over the frame and by moving these particles with person activity we extract features. These features help us to recognize person action by Adaboost classifier. Performance of the method is evaluated on a Weizmann dataset on three activities.

The advantages of the proposed method are, it does not require a tracking algorithm which makes a high time complexity. In addition, it works without detecting the person during the behavior. In future work we hope by considering overlap of the segments, some activities which are not covered in a segment completely can be recognize with higher accuracy.

## References

[1] Weinland, Daniel, Remi Ronfard, and Edmond Boyer. "A survey of vision-based methods for action representation, segmentation and recognition." Computer Vision and Image Understanding 115, no. 2 (2011): 224-241.
[2] Poppe, Ronald. "A survey on vision-based human action recognition." Image and vision computing 28, no. 6 (2010): 976-990.
[3] F. Farmani, M. Parvizimosaed, H. Monsef, A. Rahimi-Kian, A conceptual model of a smart energy management system for a residential building equipped with CCHP system, In International Journal of Electrical Power & Energy Systems, Volume 95, 2018, Pages 523-536.
[4] M. Parvizimosaed, F. Farmani, H. Monsef, A. Rahimi-Kian, A multi-stage Smart Energy Management System under multiple uncertainties: A data mining approach, In Renewable Energy, Volume 102, Part A, 2017, Pages 178-189
[5] Panagiotakis, Costas, Ilias Grinias, and Georgios Tziritas. "Automatic human motion analysis and action recognition in athletics videos." In 14th European Signal Processing Conf. 2006.
[6] S. M. R. Tousi and S. Aznavi, “Performance optimization of a STATCOM based on cascaded multi-level converter topology using multi-objective Genetic Algorithm,” in Electrical Engineering (ICEE), 2015 23rd Iranian Conference on, 2015, pp. 1688–1693.
[7] A. Ghasemkhani and Lei Yang, "Reinforcement Learning based Pricing for Demand Response," 2018 IEEE International Conference on Communications Workshops (ICC Workshops), 2018.
[8] Li, Chang-Lin, Zong-Bo Hao, and Jing-Jing Li. "Abnormal behavior detection using a novel behavior representation." In Apperceiving Computing and Intelligence Analysis (ICACIA), 2010 International Conference on, pp. 331-336. IEEE, 2010.
[9] M. A. Baferani, M. R. Chalaki, N. Fahimi, A. A. Shayegani and K. Niayesh, "A novel arrangement for improving three phase saturated-core fault current limiter (SCFCL)," 2018 IEEE Texas Power and Energy Conference (TPEC), College Station, TX, USA, 2018, pp. 1-6.

[10] B. K. P. Horn and B. G. Schunck, “Determining optical flow,” Artif.Intell., vol. 17, pp. 185–203, 1981.

[11] Amini, Mahraz, and Mads Almassalkhi. "Trading off robustness and performance in receding horizon control with uncertain energy resources." Power Systems Computation Conference (PSCC), 2018 IEEE. IEEE, 2018.

[12] Amini, Mahraz, and Mads Almassalkhi. "Investigating delays in frequency-dependent load control." Innovative Smart Grid Technologies-Asia (ISGT-Asia), 2016 IEEE. IEEE, 2016, pp. 448–453.

[13] S. M. R. Tousi and S. Aznavi, “Performance optimization of a STATCOM based on cascaded multi-level converter topology using multi-objective Genetic Algorithm,” in Electrical Engineering (ICEE), 2015 23rd Iranian Conference on, 2015, pp. 1688–1693.

[14] Younesian, D., Jafari, A.A. and Serajian, R., 2011. Effects of the bogie and body inertia on the nonlinear wheel-set hunting recognized by the hopf bifurcation theory. Int J Auto Engng, 3(4), pp.186-196.

[15] Espinosa, L.A.D., Almassalkhi, M., Hines, P., Heydari, S. and Frolik, J., 2017, March. Towards a macromodel for Packetized Energy Management of resistive water heaters. In Information Sciences and Systems (CISS), 2017 51st Annual Conference on (pp. 1-6). IEEE.

[16] Fathi, Alireza, and Greg Mori. "Action recognition by learning mid-level motion features." In Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on, pp. 1-8. IEEE, 2008.

[17] M. H. Imani, K. Yousefpour, M. J. Ghadi and M. T. Andani, "Simultaneous presence of wind farm and V2G in security constrained unit commitment problem considering uncertainty of wind generation," 2018 IEEE Texas Power and Energy Conference (TPEC), College Station, TX, USA, 2018, pp. 1-6.

[18] M. H. Imani, M. Y. Talouki, P. Niknejad and K. Yousefpour, "Running direct load control demand response program in microgrid by considering optimal position of storage unit," 2018 IEEE Texas Power and Energy Conference (TPEC), College Station, TX, USA, 2018, pp. 1-6.

[19] Schuldt, C., Laptev, I., and Caputo, B., “Recognizing Human Actions: A Local SVM Approach,” in International Conference on Pattern recognition (ICPR) Cambridge, United Kingdom, pp. 32-36, 2004.

[20] N. Fahimi, M. R. Chalaki, M. A. Baferani, M. R. Tajik and A. A. Shayegani, "Investigating the failures of defected silicon rubber insulators in salt-fog chamber," 2018 IEEE Texas Power and Energy Conference (TPEC), College Station, TX, USA, 2018, pp. 1-6.

[21] Ikizler, Nazlı, and Pınar Duygulu. "Histogram of oriented rectangles: A new pose descriptor for human action recognition." Image and Vision Computing 27, no. 10 (2009): 1515-1526.

[22] A. F. Bastani, Z. Ahmadi, and D. Damircheli, "A radial basis collocation method for pricing American options under regime-switching jump-diffusion models," Applied Numerical Mathematics, vol. 65, pp. 79-90, 2013.

[23] A. F. Bastani and D. Damircheli, "An adaptive algorithm for solving stochastic multi-point boundary value problems," Numerical Algorithms, vol. 74, pp. 1119- 1143, 2017.

[24] Pazzani, M., Feghahati, A., Shelton, C. and Seitz, A., 2018. Explaining Contrasting Categories.

[25] S. Aznavy, A. Deihimi, “Determining Switching Angles And Dc Voltages In Cascaded Multilevel Inverter For Optimizing Thd In Different Output Voltage Levels Using Genetic Algorithm”, in 6th International Conference on Technical and Physical Problems of Power Engineering, pp. 449-453, September 2010.

[26] Mohammadi, S. and Serajian, R., 2015. Effects of the change in auto coupler parameters on in-train longitudinal forces during brake application. Mechanics & Industry, 16(2), p.205.

[27] Serajian, R., 2013. Parameters’ changing influence with different lateral stiffnesses on nonlinear analysis of hunting behavior of a bogie. Journal of Measurements in Engineering, 1(4), pp.195-206.

[28] Reza Serajian, Saeed Mohammadi & Asghar Nasr (2018) Influence of train length on in-train longitudinal forces during brake application, Vehicle System Dynamics, DOI: 10.1080/00423114.2018.1456667.

[29] Espinosa, L.A.D., Almassalkhi, M., Hines, P., Heydari, S. and Frolik, J., 2017, March. Towards a macromodel for Packetized Energy Management of resistive water heaters. In Information Sciences and Systems (CISS), 2017 51st Annual Conference on (pp. 1-6). IEEE.

[30] M. Ghaisi, M. Rahmani, P. Gharghabi, A. Zoghi, and S. H. Hosseinian, “Scheduling a Wind Hydro-Pumped-Storage Unit Considering the Economical Optimization,” Am. J. Electr. Electron. Eng. , vol. 5, no. 1, pp. 16–22, 2017.

[31] M. G. Rad, P. Gharghabi, M. Rahmani, and B. Falahati, “System Vulnerability Analysis Using Graph Pathfinding Strategies in Partitioned Networks,” J. Power Energy Eng., vol. 5, no. 4, pp. 15–24, Apr. 2017.

[32] Blank, M., Gorelick, L., Shechtman, E., Irani, M. and Basri, R., 2005, October. Actions as space-time shapes. In Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on (Vol. 2, pp. 1395-1402). IEEE.