

January 2014

THE HUMAN ACTIVITY DETECTION FROM VIDEO SURVEILLANCE BASED ON HMM FOR FUTURE DEVELOPMENT IN COMPUTER ENVIRONMENT

KANCHAN GAIKWAD

Department of Computer, MGM College of Engineering, Kamothe, Panvel., kanchan.gawande@gmail.com

VAIBHAV NARAWADE

Department of Information Technology, PVPPSCOE, Sion, Mumbai, vnarawade@gmail.com

Follow this and additional works at: <https://www.interscience.in/ijcsi>



Part of the [Computer Engineering Commons](#), [Information Security Commons](#), and the [Systems and Communications Commons](#)

Recommended Citation

GAIKWAD, KANCHAN and NARAWADE, VAIBHAV (2014) "THE HUMAN ACTIVITY DETECTION FROM VIDEO SURVEILLANCE BASED ON HMM FOR FUTURE DEVELOPMENT IN COMPUTER ENVIRONMENT," *International Journal of Computer Science and Informatics*: Vol. 3 : Iss. 3 , Article 12.

DOI: 10.47893/IJCSI.2014.1147

Available at: <https://www.interscience.in/ijcsi/vol3/iss3/12>

This Article is brought to you for free and open access by the Interscience Journals at Interscience Research Network. It has been accepted for inclusion in International Journal of Computer Science and Informatics by an authorized editor of Interscience Research Network. For more information, please contact sritampatnaik@gmail.com.

THE HUMAN ACTIVITY DETECTION FROM VIDEO SURVEILLANCE BASED ON HMM FOR FUTURE DEVELOPMENT IN COMPUTER ENVIRONMENT

KANCHAN GAIKWAD¹ & VAIBHAV NARAWADE²

¹Department of Computer, MGM College of Engineering, Kamothe, Panvel.

²Department of Information Technology, PVPPSCOE, Sion, Mumbai
Email:kanchan.gawande@gmail.com, vnarawade@gmail.com

Abstract: Visual investigation of human activities related to the detection, tracking and recognition of people, and, more generally, the perceptiveness of human activities, from image sequences. Recognizing human activities from image sequences is an active area of research in computer vision. Human activity recognition (HAR) research has been on the rise because of the rapid technological development of the image-capturing software and hardware. In this paper, we propose a new approach for human action recognition from video. Here we use the HMM algorithm for recognition of activity from video. The result of this method is good as compare to other method. It consume less time as compare with other method. This result can be useful in future development as an automatic human machine interaction.

Keywords: Human activity recognition; HMM thresholding; Gaussian filtering; Content Based Video Analysis ; Interactive Applications and Environments

SECTION 1 INTRODUCTION

As an active research topic in computer visions are the dynamic scenes detection, classifying object, tracking and recognizing activity and description of behavior. Visual surveillance strategies have long been in use together information and to monitor people, events and activities. Video surveillance works as to detect moving object [1], [2], [3], [4], classify [5], [6] the detected object track [7], them through the sequence of images and analysis the behaviors. Visual surveillance technologies [8], CCD cameras, thermal cameras and night vision device are the three most widely used devices in the visual surveillance market. The main goal of visual surveillance is not only to monitor, but also to automate the entire surveillance task. The goal of visual surveillance is to develop intelligent visual surveillance to replace the traditional passive video surveillance that is proving in effective as the numbers of cameras exceed the capability of human operators to monitor them. The automated surveillance systems can be implemented for both offline like storing the video sequence and to analyses the information in that sequence. But now days online surveillance system is very much needful in all public and private sectors due to predict and avoid unwanted movements, terrorist activities in those areas. It is helpful for traffic monitoring, transport networks, traffic flow analysis, understanding of human activity [11], [10], [9], home nursing, monitoring of endangered species, and observation of people and vehicles within a busy environment along many others to prevent theft and robbery. Some of the areas where video surveillance system place a major role in many application are 1) for military security

2)patrolling of country borders 3)extracting statistics for sport activities 4)surveillance of forests for fire detection 5) patrolling of highways and railway for accident detection.

The Proposed system gives human activity detection from online video surveillance and detects single human activity from video sequence. It is useful in many other applications. The paper is arranged as section1 is including introduction. The review of methods are include in section 2.The proposed system is describe in section 3. Advantages of proposed system are in section 4. The conclusion in section 5 and references are in section 6.

SECTION 2 REVIEW OF METHODOLOGIES FOR IMAGE SEGMENTATION

Surrounding modeling is also known as Background modeling. It is currently used to detect moving objects in video acquired from static cameras. Numerous statistical methods have been developed over the recent years. The aim of this paper is to provide an extended and updated survey of the recent [12], [13], [14] researches which concern statistical background modeling. Murshed, M was proposed an edge segment based statistical background modeling [1] algorithm and a moving edge detection framework for the detection of moving objects. This paper actually focused about various methods of background modeling like traditional pixel based, edge pixel based and edge segment based approaches. He proposed this background modeling for natural image sequence with presence of illumination variation and noise. Yun Chu Zhang was analyses the background mechanism using GMM [3] model. Here

this model updates new strategy which weighs the model adaptability and motion segmentation accuracy. But these works not focus dynamic moves in frame sequence. Wei Zhou [5] was proposed the dynamic background subtraction using spatial colour binary patterns. In addition to a refined model is designed to refine contour of moving objects. This method improves the accuracy of subtracting and detecting moving objects in dynamic scenes with presence of data driven model. Richard J.Radke [14] has written survey about image change detection algorithm with several challenge issues with solved by Stauffer and Grimson's background modeling with real times image. ViBe: A Universal Background Subtraction Algorithm [6] for Video Sequences paper presence a technique for motion detection which stores a set of value taken in the past in the same location or in the neighborhood. It then compares this set to current pixel value in order to determine whether the pixel belongs to the background and to adopt the model which substitutes from the background model.

**SECTION 3
HUMAN ACTIVITY RECOGNITION FROM VIDEO**

In the proposed system live video is taking as input and then image segmentation is done by using Gaussian mixture model. The Gaussian mixture model has different methods for subtraction of background from image frame. Here we used the frame differencing method for selection of foreground image from sequences of frames. Further we used the Noise and shadow immune features of Gaussian mixture model. As shown in block diagram of proposed system as follows

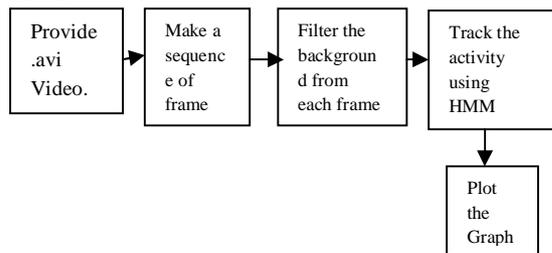


Figure 1. Shows the block diagram for human activity recognition.

This technique reduces the number of iteration for selecting human features from Sequences of frames. It means that we start by extracting the area that contains the person performing the activity, i.e., region of interest (ROI). Figure 2 shows an example result of background subtraction. A rectangular ROI is obtained from the result of background subtraction after noise and shadow removal as shown in Figure 2.b.



Figure 2. a) A frame from a aerobic sequence, b) The ROI obtained after background subtraction.

The background subtraction equation becomes

$$B(x; y; t) = I(x; y; t - 1) \\ |I(x; y; t) - I(x; y; t - 1)| > Th \quad (1)$$

$$K = \underset{K}{\operatorname{argmin}} \left\{ -\sum_{i=0}^n \ln f(x_i | \theta(K)) + \frac{K-1}{2} \ln(n) + \frac{d^2 + 3d}{4} K \ln(n) \right\} \quad (2)$$

After we get the estimated model order, we can model each type of activity by Gaussian Mixture in the feature space. Then, the ROI is partitioned into 64 blocks, $B(k)$, with equal sizes, where $k = 1, \dots, 64$. The average optical flow vector for every block is then computed by:

$$\bar{O}_k = \begin{bmatrix} \bar{O}_{kx} \\ \bar{O}_{ky} \end{bmatrix} = \frac{1}{n} \sum_{i,j \in B(k)} \begin{bmatrix} \bar{o}_x(i,j) \\ \bar{o}_y(i,j) \end{bmatrix} \quad (3)$$

where n is the number of pixels in a single block. Then, we compose the vector

$$O = [\bar{O}_1, \bar{O}_2, \dots, \bar{O}_{64}]^T$$

every frame to represent its motion feature vector, where each element contains two components for the x and the y directions.

As we get the foreground image now the selection of features like position and speed of object. The selection of position object is done by using the simple Average Theorem and combining velocity estimate over different time interval algorithm is used for selecting the speed of object.

Once the ROI (Region of Interest) is get then human activity is tracking from video using HMM thresholding algorithm. In this algorithm the maximum value of distance between frame sequences is considered as motion of object. We used both X and Y to denote video clips, and in the conditional probability expressions they represent the corresponding feature vector sequences. Based on this idea, the recognition result can be obtained as follows:

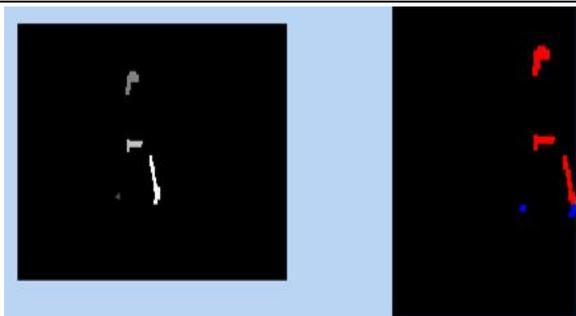


Figure 3. Tracking of activity from Region of Interest.

For recognition of activity we calculate the deviation. It is calculated for each matrix. In mat lab code The deviation is calculated as:

```
Deviation = sum (hsvCompare (pixels
(:, :, f), motionMean, shadowLevel) ./ motionDeviation,
3);
```

It considers the maximum distance between two frames as HSVCompare. Using the deviation the Sigma values is calculated in matlab as

```
sigma(:, :, 1) = sigma(:, :, 1) / (2 * pi);
```

It is unique value for each activity the sigma value is calculated for each group of matrix and calculated sigma is get compare with the trained HMM for activity detection. The recognized activity is plot using graph by change of position of with different speed of human as

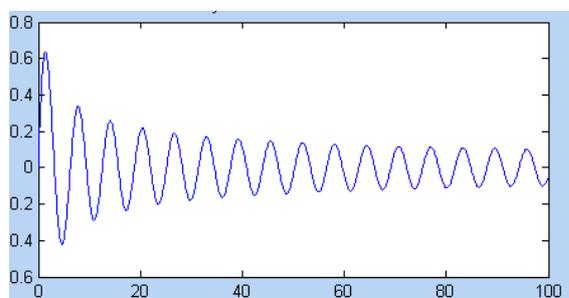


Figure 4. Graph of recognized activity position by speed of human activity.

SECTION 4 APPLICATION OF HAR

In this section, we are focusing on a few application areas that will highly the potential impact of vision-based activity recognition systems.

1) Behavioral Biometrics: It is based on physical or behavioral cues of Human. The conventional approaches are based on fingerprint, face or iris. In recent times, 'Behavioral Biometrics' have been raises fast popularity, The advantage of this approach is that subject assistance is not necessary and it can carry on without interrupting or interfering with the subject's activity. Since observing performance

implies longer term observation of the subject, approaches for action-recognition extend naturally to this task. Currently, the recent example of behavioral biometric is human gait [15].

2) Content Based Video Analysis: Now a day's video has become a part of our everyday life. With video sharing websites experiencing persistent growth, it has become necessary to develop professional indexing and storage schemes to develop user experience. This requires knowledge of patterns from raw video and summarizing a video based on its content. Content-based video summarization has been getting improved interest with corresponding advances in content-based image retrieval (CBIR) [16]. Summarization and retrieval of consumer content such as sports videos is one of the most commercially workable applications of this technology [17].

3) Security and Surveillance: Security and surveillance systems have usually depend on a network of video cameras monitored by a human operator who needs to be conscious of the activity in the camera's field of view. As recent growth in the number of cameras and deployments, the effectiveness and correctness of human operators has been increase. Hence, more security agencies are looking for vision-based solutions for security tasks which can replace by human worker. Automatic detection of anomalies in a camera's field of view is one such problem that has attracted attention from vision researchers [18]. Such more application involves searching for an activity of interest in a huge database by learning patterns of activity from long videos [19], [20].

4) Interactive Applications and Environments: To know the interface between a computer and a human remains one of the continuing challenges in designing human computer interfaces. Visual cues are the most key mode of nonverbal contact. Effective utilization of this mode such as gestures and activity holds the assure of helping in creating computers that can better interact with humans. Similarly, interactive environments such as smart rooms [21] that can react to a user's gestures can benefit from vision based methods.

5) Animation and Synthesis: The entertainment, gaming and animation industry depend on synthesizing sensible humans and human motion. The glamorous industry on the other hand has conventionally dependent more on human animators to provide best quality animation. However, this trend is rapid altering [22]. The new development in algorithms and hardware, much more realistic motion synthesis is now possible. A associated application is learning in replicated environments. Examples of such applications are training of military soldiers,

fire-fighters and other rescue personnel in hazardous situations with simulated subjects.

**SECTION 5
CONCLUSION AND ANALYSIS OF RESULT**

In this paper we use Gaussian mixture model for subtracting background from image frame sequence to get the area where the motion of human is present. This is the selection of foreground image from frame. The feature selection of image is done by simple Average Theorem, combining velocity estimate over different time interval algorithm. Then using HMM we track and recognizes the motion of human from image sequence frame.

The proposed method gives best result because; it uses the special feature of Gaussian mixture model to immune the noise and shadow from the image frame. It minimizes the no of iteration for selecting the foreground image and most promising features of object. The trained HMM track and recognized the activity of human. The proposed method uses the frames of video as shown in following figure



Figure 5. Shows the frames of video.

The confusion matrix of this method is as follows:

Actions	Walki ng	Runni ng	Gymnasi um work	Standi ng
walking	98	1	-	1
Running	1	98	-	1
Gymnasi um Work	1	1	97	1

Standing	1	1	-	98
----------	---	---	---	----

Table1. Confusion matrix for activity recognition in(%)

Sr.no	Method	Recognition Rate	Features considered
01	Human Activity Recognition Based on Spatial Transform in Video Surveillance	Walking- 88% Kick- 94% Sit down- 98%	Area, angle, centroid
02	HMM-based Human Action Recognition Using Multiview Image	Accuracy rate is 87%	silhouette, optical flow, and combined features
03	A Comparison of HMMs and Dynamic Bayesian Networks for Recognizing Office Activities	Using DBN 96.7% and using hmm 94.2%	Audio, Video, Keyboard and Mouse
04	Recognition of human activities using layered hidden markov models	Recognition rate 80%	Left and right hand movement
05	Proposed System	Walking 98% Running 98% Gymnasium work out 97% Standing 98% Over all recognition rate 98%	Position and Speed of object

The recognition rate of this method is good as compare with the other methods. This method require the less time for the recognition of activity of human. The error rate is reduces to 1.2%. The recognition rate of activity is increases and within less time. But this method is useful for the system which requires only the signals of recognized activity like instruction for robots. The comparison is shows as in Table 2.

Table 2. Comparison between proposed method with other methods.

The problem faced by the proposed system is the angel of view. It get trouble to detection of activity from video which having front view.

**SECTION 6
REFERENCES:**

- [1] M. Murshed, A.Ramirez, O.Chae, (2010) "Statistical Background Modelling: An Edge Segment Based Moving Object Detection Approach" IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS).
- [2] Peng Chena, Xiang Chenb, Beibei Jina, Xiangbing Zhua (2012) "Online EM Algorithm for Background Subtraction".
- [3] Lun Zhang, Li S.Z, Xiaotong Yuan, Shiming Xiang, (2007) "Real-time Object Classification in Video Surveillance Based on Appearance Learning" IEEE Conference on Computer Vision and Pattern Recognition.
- [4] R.N Hota, V.Venkoparao, A.Rajagopal, (2007) "Shape Based Object Classification for Automated Video Surveillance with Feature Selection", 10th International Conference on Information technology, ICIT.
- [5] Hong Lu, Hong Sheng Li, Lin Chai, Shu Min Fei, Guang Yun Liu, (2011) "Multi-Feature Fusion Based Object

- Detecting and Tracking”, journal on Applied Mechanics and Materials.
- [6] C. Beyan, A Temizel, (2012), “Adaptive mean-shift for automated multi object tracking” IET on Computer Vision.
- [7] Blanco Adán, Carlos Roberto del and Jaureguizar Nuñez, Fernando and García Santos,(2011) “Bayesian Visual Surveillance, a Model for Detecting and Tracking a variable number of moving objects”, IEEE International Conference on Image Processing, 13-14 de Enero del 2011.
- [8] Arun Hampapayr, Lisa Brown, Jonathan Connell, Ahmet Ekin, Norman Haas,(2005) “Smart Video Surveillance Exploring the concept of multiscale spatiotemporal tracking”, IEEE Signal Processing magazine, pp. 38-50.
- [9] NeilM Robertson and IanD Reid, (2011) “Automatic Reasoning about Causal Events in Surveillance Video”, EURASIP Journal on Image and Video Processing .
- [10] Jenq-Neng Hwang, (2011) “Keynote talk 4: Automated understanding of video object events in a distributed smart camera network”, International Conference on Advanced Technologies for Communications (ATC).
- [11] Ali, S.F.; Jaffar, J.; Malik, A.S.,(2012) “Proposed framework of Intelligent Video Automatic Target Recognition System (IVATRs)”, National Postgraduate Conference (NPC).
- [12] S. Radke, O.Andra, Al-Kofahi, and B. Roysam,(2005) “Image change detection algorithms: A systematic survey,” IEEE Trans. Image Process., vol. 14, pp. 294–307.
- [13] Y. Benezeth, P. Jodoin, B. Emile, H. Laurent, and C. Rosenberger,(2008) “Review and evaluation of commonly-implemented background subtraction algorithms,” in Proc. IEEE Int. Conf. Pattern Recognit., pp. 1–4.
- [14] S. Elhabian, K. El-Sayed, and S. Ahmed, (2008) “Moving object detection in spatial domain using background removal techniques—State-of-art,” Recent Pat. Computer. Sci., vol. 1, pp. 32–54.
- [15] S. Sarkar, P. J. Phillips, Z. Liu, I. R. Vega, P. Grother, and K. W. Bowyer, “The HumanID gait challenge problem: Data sets, performance, and analysis,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27, no. 2, pp. 162–177, 2005.
- [16] Y. Rui, T. S. Huang, and S. F. Chang, “Image retrieval: current techniques, promising directions and open issues,” Journal of Visual Communication and Image Representation, vol. 10, no. 4, pp. 39–62, 1999.
- [17] S. F. Chang, “The holy grail of content-based media analysis,” IEEE Transactions on Multimedia, vol. 9, no. 2, pp. 6–10, 2002.
- [18] H. Zhong, J. Shi, and M. Visontai, “Detecting unusual activity in video,” Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, pp. 819–826, 2004.
- [19] C. Stauffer and W. E. L. Grimson, “Learning patterns of activity using real-time tracking,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 8, pp. 747–757, 2000.
- [20] W. Hu, D. Xie, T. Tan, and S. Maybank, “Learning activity patterns using fuzzy self-organizing neural network,” IEEE Transactions on Systems, Man and Cybernetics, vol. 34, no. 3, pp. 1618–1626, 2004.
- [21] A. Pentland, “Smart rooms, smart clothes,” International Conference on Pattern Recognition, vol. 2, pp. 949–953, 1998.
- [22] D. A. Forsyth, O. Arikan, L. Ikemoto, J. O’Brien, and D. Ramanan, “Computational studies of human motion: part 1, tracking and motion synthesis,” Foundations and Trends in Computer Graphics and Vision, vol. 1, no. 2-3, pp. 77–254, 2005.
- [23] Feng Niu and Mohamed Abdel-Mottaleb” Hmm-Based Segmentation And Recognition Of Human Activities From Video Sequences” 0-7803-9332-5/05/ ©2005 IEEE

