www.ijecs.in
International Journal Of Engineering And Computer Science ISSN:2319-7242
Volume - 3 Issue -9 September, 2014 Page No. 8317-8320

Classification of Human Action in a Controlled Environment

Asmeet Bhosale, Geetanjali Kale
Department of Computer Engineering,
PICT, Pune, India,
E-mail: asmeetb10@yahoo.co.in
Department of Computer Engineering,
PICT, Pune, India
E-mail: gilkale@gmail.com

Abstract

It is essential to monitor suspicious activities in certain places. In this paper we propose to develop an intelligent system for visual surveillance which will notify or generate an alarm if a person is found doing a suspicious activity. The system will make use of computer vision concepts like human motion analysis for posture recognition which involves certain suspicious activity.

Keywords *Computer vision, Image processing, Human motion analysis.*

Introduction

Computer vision is a field which has tremendous potential to research and develop new applications that include processing and analyzing images. Various systems have been implemented with the help of computer vision concepts which have a vast number of applications. The advanced development in this field has led to more seamless interaction between machines and human beings, where these intelligent machines assist the human factor in many areas like surveillance, medical diagnostics, smart video processing [4] due to which it requires minimal human intervention.

In this paper we try to propose a system for visual surveillance which will notify or generate an alarm if a person is found doing something suspicious in a particular environment. Certain human activities are considered suspicious in certain areas. For human motion analysis many techniques have been proposed in the past which is required in many computer vision applications. Ronald Poppe [7] has presented a survey on a vision based human action recognition which discusses classification and temporal state space models and HMMs. J. K. Aggarwal et al. [5] discusses different aspects of high level processing. In [2] human motion analysis is discussed in terms of modeling and estimation phases. Applications such as Robert Bodor et al. [3] developed a system to track pedestrians and detect situations where people might be in danger, En-Wei Huang and Li-Chen Fu [6] proposed a method which uses vision based gesture recognition to control character animation, also in [10] Khai Tran et al, present a real time online system for continuous recognition of human actions.

Rest of the paper is as follows. Section 2 discusses the related work. Section 3 discusses the proposed system. And Section 4 discusses the mathematical model.

Literature Review

L. Wang et al. [12] has proposed an automatic gait recognition method for silhouette analysis during walking.

The three main areas are body structure analysis, tracking and recognition. In body structure analysis there are two types of approaches as follows: 1) Non-model based and 2) Model-based.

2D contours, stick figures and volumetric models [13] can be used to represent the human bodies. The stick figures are the simplest for m of representing a human body. The motion of the human body can be easily recognized using stick figures. In non-model based approaches 2D and 3D representation of human body structure is used. In [14] Rowley and et al. have presented an application of the Expectation-Maximization (EM) algorithm to the global analysis of articulated motion. Their approach utilizes a kinematic model to constrain the motion estimates. The EM algorithm for motion estimation using mixture models has two important steps: 1) Expectaion step (E-step): The mixture parameters are updated by using values of the motion parameters. It computes the conditional likelihood of each pixel originating from each motion model. Next step is called 2) Maximization step (M-step): Here the motion parameters are estimated. Their experimental results with data have shown the algorithms is accurate. Many methods also use model based approaches which makes use of a model which is defined before as a standard. O'Rourke and Badler [15] have developed a model-based 3D human motion analysis system. The system has four main steps: prediction, simulation, image analysis and parsing. So these four steps are followed in a particular manner to estimate the motion of body parts. But various model based approaches are faced with challenge of mapping the model parameters with the varying human images.

In case of tracking method for human motion analysis it involves matching two frames based on pixels, on their motion and some other information. There is feature based

tracking in which features are used for matching between two frames. Also a criteria needs to be set to select a unique and efficient feature. Now tracking can be done in two ways that is single view approach and multiple perspectives approach. In single view generally a single camera view is used to take images or video. Features like 2D blobs or meshes are used mostly. Pentland et al. [16] used the blob feature in their proposed system, the human body was constructed by blobs. The feature vector of a blob is formulated as (x,y,Y,U,V) consisting of a spatial (x,y) and color (Y,U,V) information. The assigning of pixels belonging to the human body to different parts blobs was done by using the log-likelihood measure.

Now tracking by using single view approach has a disadvantage that it can monitor only a limited area. To overcome this disadvantage tracking can be done using multiple cameras so that the monitoring area is increased. Many methods have used multiple cameras for the purpose of tracking. Sato et al. [17] considered a moving human as a combination of various blobs of its body parts. The blobs were matched over the images using various parameters like their average brightness, rough 3D position and area. The 3D position of a 2D blob was calculated based on various parameters. The blobs were then merged using motion parameters from different frames.

The recognition method involves recognizing human activities, behavior and actions. The main two types of approaches are template matching and state space. In template matching the captured images are compared with some stored images in a database to conclude a result. Template matching is one of the simplest approaches in human motion recognition. On the other hand, state space models define each actin or gesture as a state. It includes traversing through different states for identifying a particular motion. Bobick and Davis [18] have presented a novel technique for identifying actions. The approach is based on temporal templates. They made use of motion energy images (MEI) and motion history images (MHI).

State space models include models like the hidden markov model (HMM). HMM is used in many computer vision algorithms. Goddard [19] used composition of events linked by time intervals in addition to HMM. And each event is independent of the movement.

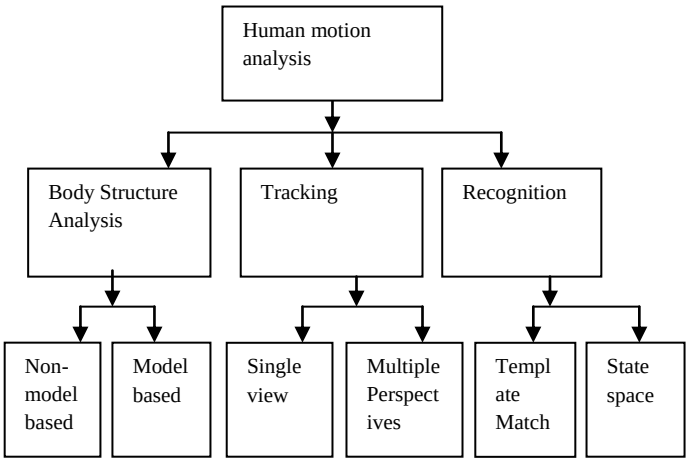

Figure 1: Human motion analysis[4].

Proposed System

The system assumes the camera is fixed which will capture the video of a person and will determine whether the activity in captured video is suspicious or not. When a suspicious activity is detected an alarm or a notification is to be generated. The system diagram in Figure 2 shows the steps through which the video needs to be processed in order to capture and detect a person's suspicious activity. In Figure 3 the system architecture diagram is shown in which the video is captured by a webcam which is given as an input to the CPU block. A database is also maintained of the templates of human suspicious activities in for various areas. The results can be displayed on a screen and alarm or other forms of outputs shall be generated too.

The system diagram of our proposed system is as follows.

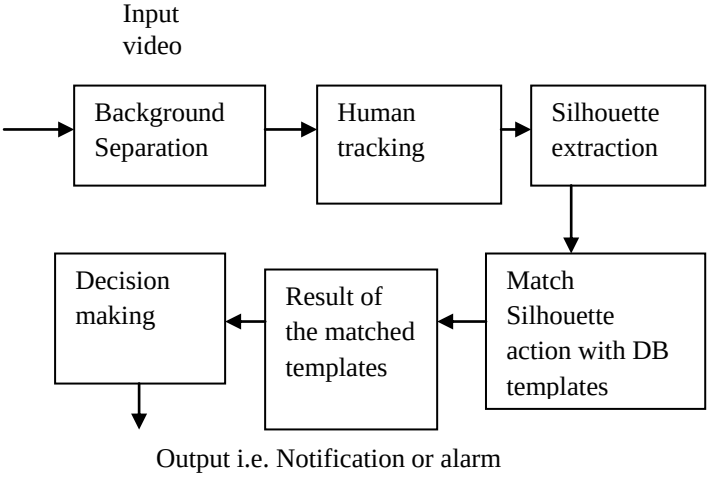

Figure 2: System diagram.

The system architecture diagram is as follows.

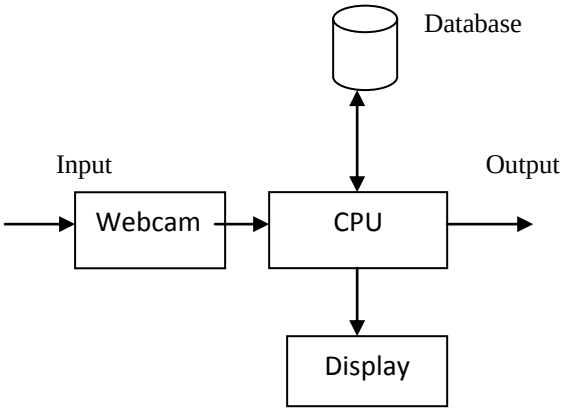

Figure 3: System Architecture diagram.

Background separation will involve having a static background and only the person doing the suspicious activity is moving. Algorithms like in Heikkila and Olli [9] can be used and also ViBe [1] is a universal background subtraction algorithm.

Next step is human identification and silhouette extraction the task will be to identify the objects of interest.

The person doing suspicious activity is identified from among objects in the scene and the silhouette of it can be obtained by applying algorithms like in [8] by Hansung Kim et al.

After obtaining the desired silhouettes we can match these silhouettes with those in the database. Figure 3 shows that a database of silhouette templates is being maintained which can help identifying the desired suspicious activity in a particular scenario. For example consider the following scenario, while operating an ATM machine a person usually operates the machine in a standing position. If the person bends for a reason for a longer duration of time it may be considered as a suspicious activity in that particular scenario. According to the areas specific templates can be maintained in the database. Algorithms like in [11] can be used for template matching.

The results of template matching will be analyzed for detection of any suspicious activity respective to a particular area. Depending on the result, if a suspicious activity is detected then an alarm or a notification will be generated accordingly.

Mathematical Model

Let S be the system to detect suspicious activity using template matching

S= {I, O, V, St, Res, P, R, S, F}

Where,

I is the input to the system

O is the output of the system

V is the set of frame images

St is the set of standard templates

Res is the result set

- Input
  I is the input set which contains the frames captured of an activity.

- Output
  O is the output set where on detection of a suspicious activity an alarm or notification will be generated.

  IR = { IR1, IR2, IR3,..... IRn} Set of intermediate results

  P = Processing function
  R = Recognition function
  Res = Final Result

- Res = $R(IR_i)$ + $Database\_Access_i$ where i=0 to n

- If Res = True Decision making

  Alarm or notification will be generated

  Else

  No alarm

Conclusion

In this paper we have proposed a new smart surveillance system for identifying suspicious activity in restricted areas. The system will have a fixed camera which will monitor a particular area. On detection of suspicious activity in the area a particular alarm or a notification will be generated.

References

[1] O. Barnich, M. V. Droogenbroeck, "ViBe: A universal Background Subtraction algorithm for video sequences", IEEE Transactions on Image Processing, 20(6):1709-1724, June 2011.

[2] R. Poppe, "Vision-based human motion analysis: An overview", ELSEVIER, Computer Vision and Image Understanding , 4-18, 2007.

[3] R. Bodor, B. Jackson, N. Papanikolopoulos, "Vision-Based Human Tracking and Activity Recognition", Citeseer, 2003.

[4] J. K. Aggarwal, Q. Cai, "Human Motion Analysis: A Review", Computer Vision and Image Understanding, Vol. 73, 428-440, 1999.

[5] J. K. Aggarwal, S. Park, "Human Motion: Modeling and Recognition of Actions and Interactions", International Symposium on 3D data processing, Visualization and Transmission, 2004.

[6] E.W. Huang, L. C. Fu, "Segmented Gesture Recognition for Controlling Character Animation", VRST, 2008.

[7] R. Poppe, "A survey on vision-baed human action recognition", ELSEVIER, Image and Vision Computing 976-990, 2009.

[8] H. Kim, R. Sakamoto, I. Kitahara, T. Toriyama, K. Kogure, "Robust Silhouette Extraction Technique Using Background Subtraction", MIRU2007, 2007.

[9] J. Heikkila, O. Silven, "A Real Time System for Monitoring Cyclists and Pedestrians", Citeseer, 1999.

[10] K. Tran, I. A. Kakadiaris, S. K. Shah, "Fusion of Human Posture Features For Continuous Action Recognition", Workshop on Sign, Gesture and Activity, 2010.

[11] D. Mohr, G. Zachmann, "Silhouette Area Based Similarity Measure for Template Matching in Constant Time", Springer, 2010.

[12] L. Wang, T. Tan, H. Ning, W. Hu, "Silhouette Analysis-Based Gait Recognition for Human Recognition", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 25, December 2003.

[13] J. K. Aggarwal, Q. Cai,W. Liao, and B. Sabata, "Non-rigid motion analysis: Articulated & elastic motion", CVGIP: Image Understanding, 1997.

[14] H. A. Rowley and J. M. Rehg, "Analyzing articulated motion using expectation-maximization", in Proc. of Intl. Conf. on Pattern Recognition, Puerto Rico, 1997, pp. 935–941.

[15] J. O'Rourke and N. I. Badler, "Model-based image analysis of human motion using constraint propagation", IEEE Trans. PAMI, **2**:522–536, 1980.

[16] C. Wren, A. Azarbayejani, T. Darrel, and A. Pentland, Pfinder: "Real-time tracking of the human body", in Proc. SPIE, Bellingham, WA, 1995.

[17] K. Sato, T. Maeda, H. Kato, and S. Inokuchi, "CAD-based object tracking with distributed monocular camera for security monitoring", in Proc. 2nd CAD-Based Vision Workshop, Champion, PA, Feburary 1994, pp. 291–297.

[18] A. F. Bobick and J. Davis, "Real-time recognition of activity using temporal templates", in Proc. of IEEE Computer Society Workshop Applications on Computer Vision, Sarasota, FL, 1996, pp. 39–42.

[19] N. H. Goddard, "Incremental Model based discrimination of articulated movement from motion features", in Proc. Of IEEE Computer Society Workshop on Motion of Non-Rigid and Articulated Objects, Austin, TX, 1994, pp. 89-95.

[20] R. T. Collins, R. Gross, J. Shi, "Silhouette-based Human Identification from Body Shape and Gait", IEEE, 2002.