

Video Surveillance in Public Transport Areas using Semantic Based Approach

Rucha D. Pathari¹, Prof. Sachin Bojewar²

P.G. Scholar,

Department of Computer Engineering,

ARMIET, Mumbai University¹

Associate Professor, Department of Computer Engineering, ,

VIT, Mumbai University²

Abstract: A complete semantics-based behaviour recognition approach that depends on object tracking has been introduced in this project. The proposed framework obtains 3-D object level information by detecting and tracking people in the scene using a blob matching technique. Based on the temporal properties of these blobs, behaviours and events are semantically recognized by employing object and inter object motion features. A number of types of behavior videos that are relevant to security in public transport areas have been selected to demonstrate the capabilities of this approach. Examples of these are abandoned and stolen objects, fighting, fainting, and loitering. Using some videos, the experimental results presented here demonstrate the performance this approach.

Keywords: Video surveillance system, Semantic based approach, Behavior classification .

INTRODUCTION

Police and security staffs are dependent on video surveillance systems to facilitate their work. These systems plays more important role in large public transportation areas such as metro stations and airports. The function of an automated surveillance system is to draw the attention of monitoring personnel to the occurrence of a user-defined suspicious behavior when it happens. Two challenges mainly come while developing fully automated behavior recognition. First, objects of interest such as people and luggage in a scene which must be found robustly, classified and tracked through time. Second, a stable means of describing events must be found.

The majority of researchers till now have invoked machine learning to detect suspicious behavior. The system proposed here is a complete semantics-based solution to the behavior detection problem that addresses the whole process from pixel to behavior level.

In contrast, the semantic approach replaces this need for training with a more straightforward process based on human reasoning and logic. This is a more feasible and viable method. This project assumes that foreground blobs are extracted in each frame using a conventional background subtraction method.

AUTOMATED SURVEILLANCE SYSTEM

Surveillance system is the process of locating moving objects over time using a camera. It has a variety of uses, some of them are: human-computer interaction, security and surveillance, video communication and compression, augmented reality, traffic control, medical imaging and video editing. It can be a time consuming process due to the amount of data that is contained in video. So its need of using object recognition techniques for tracking makes the process of video tracking more complex.

To perform the system, an algorithm analyzes sequential video frames and tracks the movement of targets between the frames. There are two major components of a visual tracking system: target representation and localization, as well as filtering and data association. Target representation and localization is mostly a bottom-up process. Locating and tracking the target object successfully is dependent on the algorithm.

EXISTING SYSTEM

Behaviour recognition is a broad term that covers a number of categories of activities, which require different means of detection to describe suspicious behaviors and activities in public transportation systems. Most of the existing systems focus on detecting a single type of behavior rather than providing a generic framework. For example, abandoned luggage detection is usually handled using low-level background subtraction methods. These methods are useful for detecting stationary foreground objects but hardly so for other complex types of behavior such as loitering or fighting, which requires experimentation and fine-tuning. This is because such behaviors manifest a broad range of variations and are very difficult to model, even if based on reasoning.

PROPOSED SYSTEM

The proposed system focuses on detecting suspicious behavior in public transportation systems. First, the proposed framework obtains 3-D object-level information by detecting and tracking people and luggage in the scene using a blob matching technique. Based on the temporal properties of these blobs, behaviors and events are recognized by employing object and inter-object motion features. A number of types of behavior that are relevant

to security in public transport areas have been selected to demonstrate the capabilities of this approach.

The framework performs object tracking in each frame, and then it goes for behavior recognition for every frame. The operation is followed by background subtraction. Further, the color histograms used seem to provide low complexity while simultaneously dealing with constantly occluding patterns. In addition to the single-object features, the inter-object features between every combination of two objects are also stored in historical sequence.

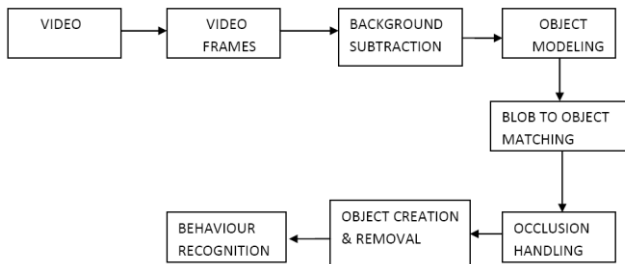


Fig.1 Proposed System

Architecture

1. Object Modelling and Blob-To-Object Matching

In object tracking approach, at each frame, a list of objects is updated by matching blobs in the current frame with objects from the previous one. This matching process is not necessarily one to-one. To match blobs and objects in two consecutive frames, color histograms and spatial information are used. The color histograms are all examined to ensure a correct update. The color histograms are adaptively updated using following equation (1).

$$\text{Histogram}_{\text{blob}, t} = \alpha \text{Histogram}_{\text{object}, t-1} + (1-\alpha) \text{Histogram}_{\text{object}, t} \quad \dots (1)$$

Where α is the learning rate.

A basic form of histogram intersection is given by the following equation (2).

$$\text{Intersection}(\text{Hist}_1, \text{Hist}_2) = \sum_{i=1}^n \min(Hist_{1,i}, Hist_{2,i}) \quad \dots (2)$$

Here, normalize the intersection by the largest of the histograms. The formula for histogram intersection then becomes:

$$\text{Intersection}(\text{Hist}_1, \text{Hist}_2) = \frac{\sum_{i=1}^n \min(Hist_{1,i}, Hist_{2,i})}{\max(\|Hist_1\|, \|Hist_2\|)} \quad \dots (3)$$

2. Occlusion Handling

Occlusion handling is a critical task because it bears on the robustness of object tracking and coherence. In this framework, the issue of which objects are occluding which is completely ignored. Dummy objects are also created for the pool that exhibits the adaptive appearance model necessary for blob matching. In a nutshell, the system also

renders the phenomenon of occlusion into a split or merge problem. Merges and splits of blobs are checked before any one-to-one associations are made.

3. Object Creation and Removal

After all blob-to-object matching cases of merging, splitting, and one-to-one matches have been processed, some of the remaining blobs and objects may still remain unmatched. An unmatched blob is ideally a new object that has just appeared in the scene. Therefore, a new object is created for each unmatched blob. On the other hand, an unmatched object could also be either an object that has just left the scene or one whose blob has been falsely undetected due to some failure in background subtraction. Therefore, a grace period of a few seconds is provided to allow for the object's recovery.

4. Behavior Semantics Recognition

Following types of behavior is classified by the proposed system:

1) **Loitering:** Loitering is useful for detecting a number of public transit situations such as drug dealing. It is defined as "the presence of an individual in an area for a period of time longer than a given time threshold."

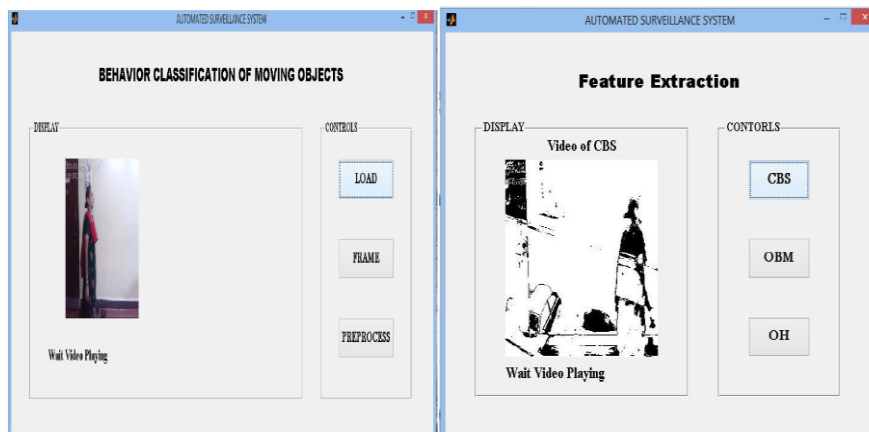
2) **Fighting:** The system indicate that the most reliable means for defining fights is in terms of the frequency of object splitting, merging or a presence of high dynamic level of activity.

3) **Meeting and Walking Together:** Although generally not considered to be suspicious, meeting and walking together may be useful in certain surveillance scenarios. This would be particularly the case were face recognition included as a feature. For an example, it might be pertinent for security purposes to flag individuals that meet with a suspicious individual.

ALGORITHM USED

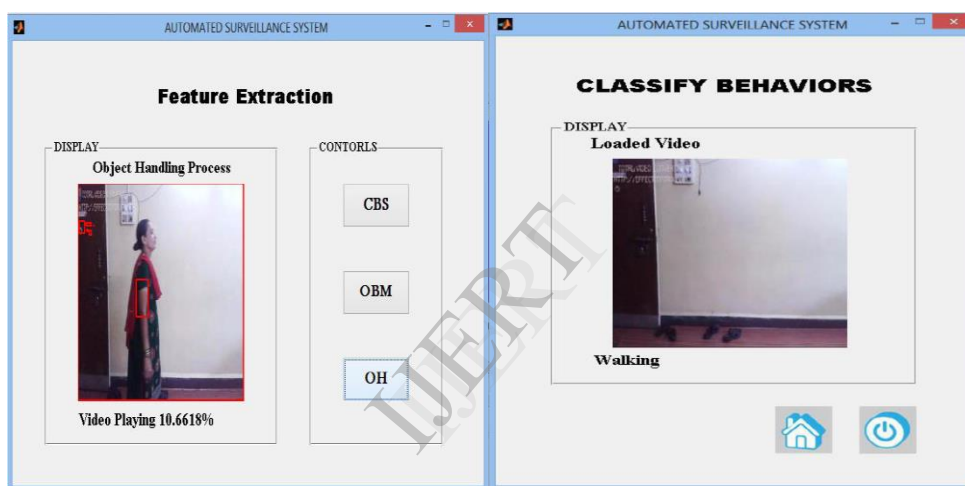
1. START
2. Load The Video.
3. Convert the Video into frames.
4. By using Gaussian filter remove noise from the frames.
5. Once the noise removal takes place, For CBS (Conventional Background Subtraction) frame differencing is used.
6. On CBS processed Frames Blob to object matching is performed by finding color histogram for each frames and checking for intersections if the TH value is greater than the specified range then the occlusion is detected, else no occlusion into the video.
7. If the occlusion is detected in the video it handles merge and splits of the objects by using Kalman Filter.
8. After merging it will check for unmatched blobs if it is found repeat step 5 to 8. the behaviors in the video loaded are classified and the message will be shown accordingly.
9. END

EXPERIMENTAL RESULTS



(a)

(b)



(c)

(d)

Fig 2. (a)Convert the video into frames (b) Background Image Subtraction(c) Object Handling Process(d)Recognize the Behavior

CONCLUSION

The proposed approaches ensure performance, adaptability, robustness against camera nonlinearities, ease of interfacing with human operators, and elimination of the training required by machine-learning-based methods. Detection of suspicious activities in public transport areas using video surveillance has attracted an increasing level of attention. In the project, suspicious behavior of the people in public area from the recorded video is identified using cluster matching. A complete semantics-based behaviour recognition approach that depends on object tracking has been introduced and extensively investigated. The approach begins by translating the objects obtained by background segmentation into semantic entities in the scene. Ultimately, behaviours are semantically defined and detected by continuously checking these records against pre defined rules and conditions.

This approach ensures performance, adaptability, robustness against clutter and camera nonlinearities, ease of interfacing with human operators, and elimination of the training required by machine-learning-based methods. Experimentation was carried out on multiple standard publicly available data sets that varied in terms of crowd density, camera angle, and illumination conditions. The experimental results demonstrated successful detection of the various activities of interest.

REFERENCES

- [1] H. Weiming, T. Tieniu, W. Liang, and S. Maybank, "A survey on visual surveillance of object motion and behaviors," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 34, no. 3, pp. 334–352, Aug. 2004.
- [2] G. L. Foresti, C. Micheloni, L. Snidaro, P. Remagnino, and T. Ellis, "Active video-based surveillance system: The low-level image and video processing techniques needed for implementation," *IEEE Signal Process. Mag.*, vol. 22, no. 2, pp. 25–37, Mar. 2005.
- [3] N. Firth, "Face recognition technology fails to find U.K. rioters," *New Scientist*, London, U.K. [Online]. Available: <http://www.newscientist.com/article/mg21128266.000-face-recognition-technology-fails-to-find-rioters.html>
- [4] L. M. Fuentes and S. A. Velastin, "Tracking-based event detection for CCTV systems," *Pattern Anal. Appl.*, vol. 7, no. 4, pp. 356–364, Dec. 2004.
- [5] M. Elhamod and M. D. Levine, "A real time semantics-based detection of suspicious activities in public scenes," in *Proc. 9th Conf. CRV, Toronto, ON, Canada, 2012*, pp. 268–275.
- [6] N. T. Siebel and S. J. Maybank, "The ADVISOR visual surveillance system," in *Proc. ECCV Workshop ACV, 2004*, pp. 103–111.
- [7] Z. Zhang, T. Tieniu, and H. Kaiqi, "An extended grammar system for learning and recognizing complex visual events," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 2, pp. 240–255, Feb. 2011.
- [8] D. Demirdjian and C. Varri, "Recognizing events with temporal random forests," in *Proc. Int. Conf. Multimodal Interfaces, Cambridge, MA, 2009*, pp. 293–296.
- [9] D. Weinland, R. Ronfard, and E. Boyer, "A survey of vision-based methods for action representation, segmentation and recognition," *Comput. Vis. Image Underst.*, vol. 115, no. 2, pp. 224–241, Feb. 2011.
- [10] M. Blank, L. Gorelick, E. Shechtman, M. Irani, and R. Basri, "Actions as space-time shapes," in *Proc. 10th IEEE ICCV, 2005*, vol. 2, pp. 1395–1402.
- [11] N. D. Bird, O. Masoud, N. P. Papanikolopoulos, and A. Isaacs, "Detection of loitering individuals in public transportation areas," *IEEE Trans. Intell. Transp. Syst.*, vol. 6, no. 2, pp. 167–177, Jun. 2005.
- [12] N. Bird, S. Atef, N. Caramelli, R. Martin, O. Masoud, and N. Papanikolopoulos, "Real time, online detection of abandoned objects in public areas," in *Proc. IEEE ICRA, 2006*, pp. 3775–3780.
- [13] L. Sijun, Z. Jian, and D. Feng, "A knowledge-based approach for detecting unattended packages in surveillance video," in *Proc. IEEE AVSS, 2006*, p. 110.
- [14] S. Blunsden and R. B. Fisher, "The BEHAVE video dataset: Ground truthed video for multi-person behavior classification," *Annu. BMVA*, vol. 2010, no. 4, pp. 1–11, 2010.
- [15] S. Blunsden, E. Andrade, and R. Fisher, "Non parametric classification of human interaction," in *Proc. 3rd Iberian Conf. Pattern Recog. Image Anal., Part II, Girona, Spain, 2007*, pp. 347–354.
- [16] P. Fatih, "Detection of temporarily static regions by processing video at different frame rates," in *Proc. IEEE Conf. AVSS, 2007*, pp. 236–241.
- [17] A. Singh, S. Sawan, M. Hanmandlu, V. K. Madasu, and B. C. Lovell, "An abandoned object detection system based on dual background segmentation," in *Proc. 6th IEEE Int. Conf. AVSS, 2009*, pp. 352–357.
- [18] A. Hakeem, Y. Sheikh, and M. Shah, "CASEE: A hierarchical event representation for the analysis of videos," in *Proc. 19th Nat. Conf. Artif. Intell., San Jose, CA, 2004*, pp. 263–268.