ISSN (Online) : 2278-1021
ISSN (Print) : 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 3, Issue 6, June 2014

Extract Semantic Context Information for Intelligent Video Surveillance of Traffic scenes

Revanasidda[1], R.P.Rajeshwari[2]

IV Semester, M.Tech., Department of Computer Science & Engineering, Rao Bahadur Y.Mahabaleswarappa Engineering College, Bellary, India[1]

M.Tech(CSE) Lecturer, Department of Computer Science & Engineering, Rao Bahadur Y.Mahabaleswarappa Engineering College, Bellary, India[2]

Abstract: Surveillance are nothing but an monitoring, here we use an video surveillance to provide Security to the people to avoid the unusual or abnormal activities in the field. Visual surveillance systems are mainly used for public security. we try to Extract here semantic context information which including object-specific context and scene-specific context information to make an intelligent system with robust object detection, tracking, warning signals, classification and abnormal event detection. In object-specific context information, objects are classified into an meaningful categories. Object classification is also called as a co trained classifier, which takes advantage of the multi view information of objects and collects the strong features to identify an object and reduces the number of labeling training samples, is learned to classify objects into pedestrians or vehicles with high object classification performance. For each kind of object, we learn its corresponding semantic features includes: motion pattern each vehicles, width distribution, paths, and entry/exist points. Based on this information, it is efficient to improve object detection, tracking, abnormal event detection and warning signals recognition.

Index Terms: object detection, object classification, Gaussian mixture model (GMM) and graph cut, object classification, object tracking, warning signals, video surveillance.

I. INTRODUCTION

Surveillance systems are mainly used for public security by using surveillance on the traffic field its very efficient to avoid unusual things in the field. There has been an increasing demand for visual surveillance systems :more and more surveillance cameras are used in public areas such as airports, banks, malls, and subway stations. However, they are not optimally used due to the manual observation of the output, which is expensive and unreliable. Automated surveillance systems aim to integrate real-time and efficient computer vision algorithms in order to assist human operators. To provide security for the public we need to solve commonly encountered surveillance problems which are of object detection, object classification, object tracking, warning signals recognition and abnormality detection. we attempt to solve these problems by extracting semantic context information.

II. LITERATURE SURVEY

A typical semi-supervised learning algorithm is the co training approach proposed by Blum and Mitchell . The basic idea is to train two classifiers on two independent “views” (features) of the same data, using a relatively small number of examples. These classifiers then go through unlabeled examples, label them, and add the most confident predictions to the labeled set of the other classifier. In other words, the classifiers train each other using the unlabeled data. Some work has proved that co training can find a very accurate classification. Inspired by the co training idea, we propose an unsupervised learning method by combining multiple features with small labeled data for training two classifiers, which are adopted to classify a foreground into a pedestrian or vehicle. The classifiers collaboratively classify the unlabeled data and use this newly labeled data to update each other. In our algorithm, classifiers are not pre-trained, and two relatively independent features are used: object-specific context features and multi block local binary pattern(MB-LBP) features as the object representation. Each feature is used to train a classifier, and their outputs are combined to give the final classification results. Experiments demonstrate that co training can generate an accurate classifier conveniently and effectively. Wanget al. use the distributions of observations (positions and moving directions of objects) on the trajectories for trajectory analysis, but do not take into account the integrity of each trajectory. Based on the definition in the continuity of trajectory is ignored. After clustering trajectories, semantic scene models are obtained for each cluster. Paths can be detected by modeling the spatial extents of trajectory clusters . Entry and exit points are detected at the ends of paths based on the velocity distribution . Makris and Ellis detect these points from start/end points of trajectories by the Gaussian mixture models (GMM).

III. EXISTING SYSTEM

Automated surveillance systems aim to integrate real-time and efficient computer vision algorithms in order to assist human operators. This is an ambitious goal which has attracted an increasing amount of researchers to solve commonly encountered surveillance problems of object detection, object classification, object tracking, and

ISSN (Online) : 2278-1021
ISSN (Print) : 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
Vol. 3, Issue 6, June 2014

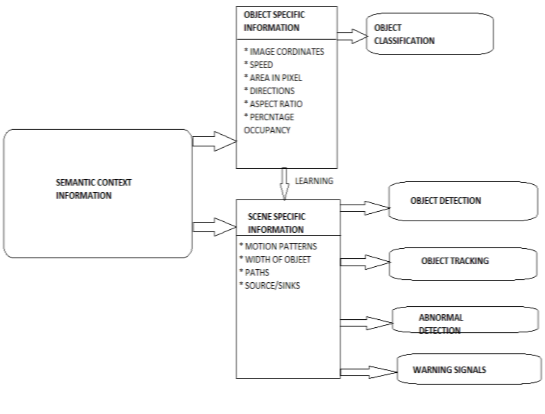
abnormality detection over the years. attempt to solve these problems by mining semantic context information. However, it is common that several objects form one big blob, which is called multi-object, because of the angle of the camera, shadow, and moving objects near each other. Since a multi-object is detected as one foreground, it is difficult to obtain the appearance feature of each single object. Therefore, it is difficult to classify and track the objects. Number of works has been proposed to solve the crowd segmentation problem, which emphasized locating individual humans in a crowd. In head detection is used to help locate the position of humans. Dong *et al.* proposed a novel example-based algorithm that maps a global shape feature by Fourier descriptors to various configurations of humans directly and use locally weighted average to interpolate for the best possible candidate configuration. In addition, they use dynamic programming to mitigate the inherent ambiguity.

Disadvantages

- The directions of motion of objects are different, their postures will change, which may cause these features not to be feasible. In addition, objects in a group may have similar color, texture, and shape features.
- The object shapes in video may hange drastically under different camera view angles. In addition, the detected shapes may be noised by shadow or other factors.

IV. PROPOSED SYSTEM ADVANTAGES

- The main goal of our method is Semantic context information includes object-specific context information and scene-specific context information. Object-specific context information coordinates area in pixels, speed, and direction of motion, aspect ratio, and percentage occupancy.
- Then, the semantic context information is adopted to improve object detection, classification and tracking, and detect abnormal events.
- The properties of objects in the scene image and can be learned from long-term observations, which can be used to distinguish objects.



V. MODULES

Object detection

object detection is main task in video surveillance, used to find and specify objects in the field. The field may contains number objects suppose there are few objects in the field its very difficult to get or track both of the object because each connected component of the foreground usually corresponds to an object; this kind of blob is denoted as single-object. However, it is common that several objects form one big blob, which is called multi-object, because of the angle of the camera, shadow, and moving objects near each other. Since a multi-object is detected as one foreground, it is difficult to obtain the appearance feature of each single object. To avoid these problems here we use an scene specific context information, which reflects the motion rules of the object including direction of motion size of the object at a certain location.

Object tracking (lane departure warning system)

object tracking and abnormal event detection are two basic task in video surveillance.
we can detect abnormal events, improve object tracking, and help to guide vehicles.
This module shows how to track and detect road lane markers in video sequence and notifies the driver if they are moving across a lane. Algorithm. The module implements this algorithm using the following steps:

1) Detect lane markers in the current video frame.
2) Match the current lane markers with those detected in the previous video frame.
3) Find the left and right lane markers.
4) Issue a warning message if the vehicle moves across either of the lane markers. To process low quality video sequences, where lane markers might be difficult to see or are hidden behind objects, the module waits for a lane marker to appear in multiple frames before it considers the marker to be valid. The module uses the same process to decide when to begin to ignore a lane marker. We learn the scene model by using the observation of the track over a long period of time, the main aim of the object tracking is to guide the vehicle with their entry and exit point in the field.

Learning Scene-Specific Context Information

Scene-specific context features reflect the properties of objects in the scene image and can be learned from long-term observations, which can be used to distinguish objects. It is time- consuming and needs a lot of storage space to obtain these features for each pixel in the scene image.

Learning Motion Patterns

There are two types of trajectories. One belongs to vehicles, and the other belongs to pedestrians. For each type of trajectory, the motion patterns of each block can be viewed as Gaussian distributions from statistic point of view. Because each block may contain many motion patterns, we adopt the multiple Gaussian models to represent them.

The scene specific context information reflects the following properties by using Gaussian mixture model algorithm.

ISSN (Online) : 2278-1021
ISSN (Print) : 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
Vol. 3, Issue 6, June 2014

VI. ALGORITHM USED(GMM)

By using gmm we are going to find. Motion pattern, width, path, source/sinks of each block.

Step1: A trajectory can be obtained by tracking the cancroids of an object and it can be written as T[{X1,Y1}{….}{Xn,Yn]. Where {Xn,Yn} is point in the scene.

Step2:each model in the scene is mixture of K gaussion distribution Tt=(at,bt,ct,vt) where at,bt,ct,vt are the parameter of the trajectory Tt

Step3: the probability of each block can be written as

$$P(T_t) = \sum_{i=1}^{K} w_{i,t} \times \eta(T_t, u_{i,t}, \Sigma_{i,t})$$

Where: $w_{i,t}$ is weight of the parameter of time t $\eta(T_t, u_{i,t}, \Sigma_{i,t})$ is ith normal distribution.

Step4:first gaussion component that matches the test trajectories are as follows

$$w_{i,t} = (1-\alpha)w_{i,t-1} + \alpha(M_{i,t})$$
$$u_{i,t} = (1-\rho)u_{i,t-1} + \rho T_t$$
$$\sigma^2_{i,t} = (1-\rho)\sigma^2_{i,(t-1)} + \rho(T_t - u_{i,t})^T(T_t - u_{i,t})$$
$$\rho = \alpha\eta(T_t | u_{i,t}, \sigma_{i,t})$$

Where: $\sigma_{i,t} = (\sigma^a_{i,t}, \sigma^b_{i,t}, \sigma^c_{i,t})^T$, $M_{i,t}$ is 1 for the model which matched and 0 for the remaining models.

Step5: width distribution of each block can be learned for a force ground width $w_{i,t}$ in block (X0,Y0) at time t.

Step6: paths are composed of blocks with similar motion patterns. We consider minimizing the following functions.

$$E(L) = \sum_{p \in S} D_p(L_p) + \sum_{(p,q) \in \mathcal{N}} V_{p,q}(L_p, L_q)$$

Where: S is the block lattice, N is the pair wise neighborhood , Dp(Lp) is cost of the block p, Vp,q(Lp,Lq) spatially neighborhood blocks.

Step7: Dp(Lp) can be calculated as

$$D_p(L_p) = \begin{cases} \frac{d_1}{d_1+d_2}, & \text{if } L_p = 1 \\ \frac{d_2}{d_1+d_2}, & \text{if } L_p = 0 \end{cases}$$

Where: d1 & d2 represents the similarities between the motion pattern of each block.

$$V_{p,q}(L_p, L_q) = \begin{cases} 0, & \text{if } L_p = L_q \\ d_0, & \text{otherwise} \end{cases}$$

Step8:

Where: $d_0,$ is the constant which represents the minimum distance between motion patterns of each block.

Object classification

Objects are classified into meaningful categories. The classifiers are trained as follows. For a certain scene, some samples are labeled to train the two classifiers, then the classifiers are used to classify unlabeled examples to obtain their labels and add those newly labeled examples which are confident enough to update the training set for each other. This learning process can be repeated many times.

Warning signals

This is the another module or option of my project, warning signals, the vehicle are going to recognize signals by automatically, from this we can avoid unusual activities in the field.

VII. CONCLUSION

Video surveillance have increasing demand many fields are using this system here we track the semantic context information for intelligent video surveillance of traffic scenes. we introduce how to learn scene-specific context information from object-specific context information. Then, object classification is improved by combining of multiple features. Based on the learned information, we adopt it to improve object detection and tracking, and detect abnormal events. From this we can avoid and improve the robust object detections.

REFFERENCE

[1] O. Javed, K. Shafique, and M. Shah, “Automated visual surveillance in realistic scenarios,” IEEE Multimedia, vol. 14, no. 1, pp. 30–39, Jan. –Mar. 2007.

[2] R. T. Collins, A. J. Lipton, and T. Kanade, “Introduction to the special section on video surveillance,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 22, no. 8, pp. 745–746, Aug. 2000.

[3] T. Cucinotta, L. Palopoli, L. Abeni, D. Faggioli, and G. Lipari, “On the integration of application level and resource level qos control for realtime applications,” IEEE Trans. Ind. Inf., vol. 57, no. 4, pp. 479–491, Nov. 2010.

[4] E. Camponogara, A. de Oliveira, and G. Lima, “Optimization-based dynamic reconfiguration of real-time schedulers with support for stochastic processor consumption,” IEEE Trans. Ind. Inf., vol. 57, no. 4, pp. 594–609, Nov. 2010.

[5] G. Wang, L. Tao, H. Di, X. Ye, and Y. Shi, “A scalable distributed architecture for intelligent vision system,” IEEE Trans. Ind. Inf., vol. 8, no. 1, pp. 91–99, Feb. 2012.