International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Impact Factor (2012): 3.358

A Novel Approach of Mining Semantic Context Information for Intelligent Video Surveillance of Traffic Scenes

M.Kameshwara Rao[1], P. Bhavya Sree[2]

[1, 2]Assistant Professor, Department of ECM, K L Univresity Greenfields, Vaddeswaram, Vijayawada, Guntur District, Andhra Pradesh 522502, India

Abstract: ***Automated visual surveillance systems are attracting extensive interest due to public security. In this paper, we attempt to mine semantic context information including object-specific context information and scene-specific context information (learned from object-specific context information) to build an intelligent system with robust object detection, tracking, and classification and abnormal event detection. By means of object-specific context information, a cotrained classifier, which takes advantage of the multi view information of objects and reduces the number of labelling 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 scene specific context information: motion pattern, width distribution, paths, and entry/exist points. Based on this information, it is efficient to improve object detection and tracking and abnormal event detection. Experimental results demonstrate the effectiveness of our semantic context features for multiple real-world traffic***

Keywords: object detection, object tracking, abnormal event detection, morphological operation Estimation and video surveillance

1. Introduction

Surveillance plays an important part in security systems for public areas such as airports, banks, malls, and subway stations. It closely monitors the actions of the specific individual. They monitors and track the movements and data. However, 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 abnormality detection over the years. In this paper, we attempt to solve these problems by mining semantic context information.

Object detection is the process of finding real-world objects such as vehicles, buildings, humans in images or videos. A stationary camera has been fixed and it is used to find the objects present in the image. If there are few objects in the scene, 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.

Object tracking the learned information, we can detect abnormal events, improve object tracking, and help guide vehicles. Recently, many approaches have been proposed to learn motion patterns. Some of them are based on trajectory analysis. These methods can be categorized into two classes: spatial distance-based methods and spatial distribution-based methods. Spatial distance- based methods take only the pair wise similarities between trajectories.

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. Adjacent pixels in the scene image have similar scene context features; therefore, it is viable to cut the scene into blocks, where is the number of rows and is the number of columns. The size of each block is relatively small, hence the motion pattern and size of a moving object in a certain block are considered to be constant.

Learning Motion Patterns classified into vehicles or pedestrians, and 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

2. Object Detection

Object detection is the most important task in video surveillance. This Technique detects a specific object based on finding point correspondences between the target image and the reference image. In the situation of stationary cameras, background modeling, is a widely used technique to extract the moving pixels (foreground). If there are few objects in the scene, each connected component of the foreground (blob) 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. Therefore, it is difficult to classify and track the objects. A number of works have been pro-posed to solve the crowd segmentation problem, which emphasized locating individual humans in a crowd.

3. LDA Based Classification

Linear Discriminant is to find linear combination of features which characterizes more than one classes or events. These methods are used in statistics, pattern recognition and machine learning. This classification is based on Analysis of Variance (ANOVA) concept and regression analysis. They express one dependent variable as a linear combination of other features. LDA and PCA (Principal component analysis) are inter related to each other. They are for combination of variables which leads to provide more details about their data. When the measurements are continuous, LDA works.

Consider a set of features, attributes, variables or measurements) for each sample of an object or event with known class y. This set of samples is called the training set. The classification problem is then to find a good predictor for the class y of any sample of the same distribution (not necessarily from the training set) given only an observation LDA approaches the problem by assuming that the conditional probability density function and are both Multivariate normal distribution with mean and covariance parameters respectively. Under this assumption, the Bayes optimal solution is to predict points as being from the second class if the log of the likelihood ratios is below some threshold T, so that;

Without any further assumptions, the resulting classifier is referred to as QDA. LDA also makes the simplifying Homoscedastic assumption (*i.e.*that the class covariances are identical, so that the covariances have full rank. In this case, several terms cancel and the above decision criterion becomes a threshold.

This means that the criterion of an input being in a class y is purely a function of this linear combination of the known observations. It is often useful to see this conclusion in geometrical terms: the criterion of an input being in a class y is purely a function of projection of multidimensional-space point onto vector (thus, we only consider its direction). In other words, the observation belongs to y if corresponding point is located on a certain side of a hyperplane is perpendicular. The location of the plane is defined by the threshold c.

4. Semantic Concepts

Semantic file systems allows the data to be addressed by their content. They raise technical design challenges as indexes of words, tags or elementary signs are to be created and constantly updated, maintained and cached for performance to offer the desired random, multi-variate access to files in addition to the underlying, mostly traditional block-based filesystem. Traditional file systems tend to impose a burden, for example when a sub-directory layout is contradicting a user's perception of where files would be stored. Tag-based interface alleviates this hierarchy problem and enables users to query for data in an intuitive fashion. The Semantic Web provides a common framework that allows data to be shared and reused across application, enterprise, and community boundaries. The main purpose of the Semantic Web is driving the evolution of the current Web by enabling users to find, share, and combine information more easily.

5. Object Tracking and Abnormal Event Detection

Object tracking and abnormal event detection are another two important tasks in video surveillance. We learn the scene model by using the observations of tracks over a long period of time. Based on the learned information, we can detect abnormal events, improve object tracking, and help guide vehicles. Recently, many approaches have been proposed to learn motion patterns. Some of them are based on trajectory analysis. These methods can be categorized into two classes: spatial distance-based methods, and spatial distribution-based methods. Spatial distance-based methods take only the pairwise similarities between trajectories. The proposed trajectory similarities or distances include Euclidean distance, Hausdorff distance and its variations, and dynamic time warping (DTW) . These approaches have several drawbacks: they lack a probabilistic explanation for abnormality detection, require the cluster number in advance, have a high computational cost, and may not well approximate the true similarity.

6. Motion Trajectory

A Trajectory is a path followed by a moving object through space as a function of time. A satellite may be the best moving object which travels around the centre of mass. The path of a planet, an asteroid or a comet forms the orbit system and they act as trajectory also. A Trajectory is a time-ordered set of states of a dynamical system. Another example of a trajectory is the path of a projectile such as thrown ball or rock. The object moves only under the influence of a uniform gravitational force field. This can be a good approximation for a rock that is thrown for short distances for example, at the surface of the moon. In this simple approximation the trajectory takes the shape of the parabola. Generally, when determining trajectories it may be necessary to account for non uniform gravitational forces, air resistance.

Derivation of the equation of motion

Assume the motion of the projectile is being measured from a Free fall frame which happens to be at (x,y)=(0,0) at t=0. The equation of motion of the projectile in this frame (by the principle of equivalence) would be y=xtan(ø). The co-ordinates of this free-fall frame, with respect to our inertial frame would be $y=-gt^2/2$. That is $y=-g(x/v_h)^2/2$. Now translating back to the inertial frame the co-ordinates of the projectile becomes $y=-(g\sec^2ø/2v_0^2)x^2 + x\tanø$.

7. Video Surveillance

Surveillance is the monitoring of the behaviour, activities, or other changing information, usually of people for the purpose of influencing, managing, directing, or protecting them. This can include observation from a distance by means of electronic equipment (such as cameras), or interception simple, relatively no or low-technology methods such as human intelligence agents and postal. The word surveillance comes from a French Language phrase for "watching over" ("sur" means "from above" and "eviller" means "to watch").

Surveillance is very useful to governments and law enforcement to maintain, recognize and monitor threats, and prevent/investigate criminal activity. With the advent of programs such as the Total Information awareness program and ADVISE. Technologies such as Narus Insight and surveillance simulation software, and laws such as the Communications Assistance for Law Enforcement Act, governments now possess an unprecedented ability to monitor the activities of their subject.

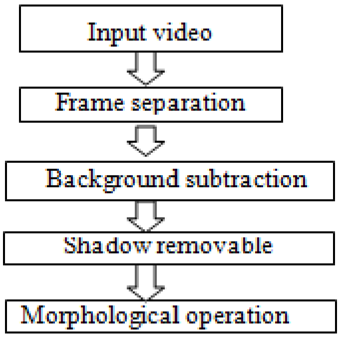


Figure 1: Block Diagram

8. Input Video

Figure 2: Input Video

9. Frame Separation

Figure 3a: Frame 2

Figure 3b: Frame 45

10. Background Subtraction

This technique is used to detect the moving object in an image. It also helps to find the foreground object and hence can be called as foreground detection. It clearly identifies the presence of an object in the image. The Intensity and the color informations plays a vital role in the background subtraction method.

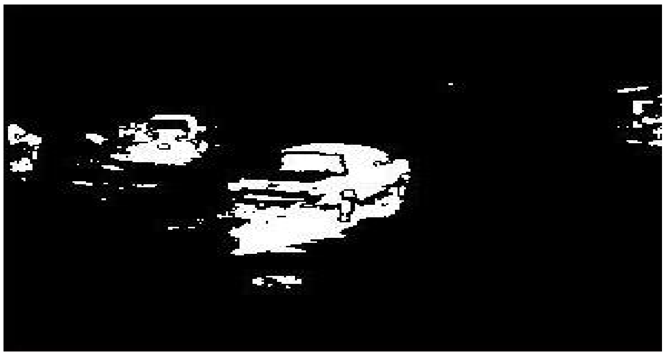


Figure 4: Background Subtraction

10.1 Shadow Removable

To get a clear image shadow should be removed from the image. In general photography, and with the advent of cameras able to capture more than 8 bits per channel, strong shadows also often characterize a high dynamic range (HDR) image. HDR images cannot always be properly displayed on current monitors. The dynamic range can be reduced if shadow present in the image is removed or attenuated.

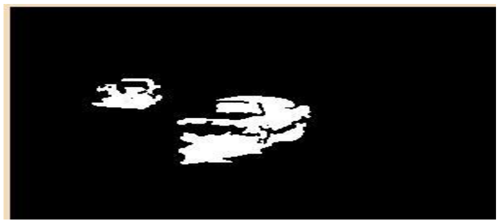
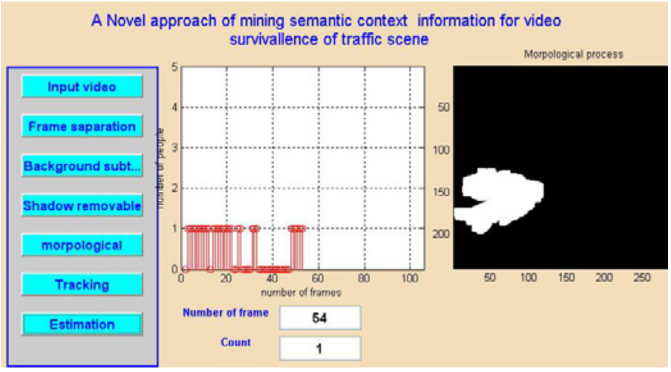


Figure 5: Shadow Removal

10.2 Estimation



11. Conclusion and Future Work

We tried with many real time videos and implemented the algorithm over it. The used videos will have many different real time traffic scenarios. We have compared our results with various similar algorithms and proved the efficiency of our algorithm. In the future, we will try for algorithm which could check for pedestrians conditions also.

Reference

[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.

[6] C. Stauffer and E. Grimson, "Learning patterns of activity using realtime tracking," IEEE Trans. Pattern Anal. Mach. Intell., vol. 22, no. 8, pp. 747–757, Aug. 2000.

[7] W. Hu, X. Xiao, Z. Fu, D. Xie, T. Tan, and S. Maybank, "A system for learning statistical motion patterns," IEEE Trans. Pattern Anal. Mach. Intell., vol. 28, no. 9, pp. 1450–1464, Sep. 2006.

[8] Saleemi, K. Shafique, andM. Shah, "Probabilistic modeling of scene dynamics for applications in visual surveillance," IEEE Trans. Pattern Anal. Mach. Intell., vol. 31, no. 8, pp. 1472–1485, Aug. 2009.

[9] T. Bcher, C. Curio, J. Edelbrunner, C. Igel, D. Kastrup, I. Leefken, G. Lorenz, A. Steinhage, and W. V. Seelen, "Image processing and behavior planning for intelligent vehicles," IEEE Trans. Ind. Electron., vol. 90, no. 50, pt. 1, pp. 62–75, Feb. 2003.

[10] J. C. Nascimento and J. S.Marques, "Performance evaluation for object detection algorithms for video surveillance," IEEE Trans. Multimedia, vol. 8, no. 4, pp. 761–774, Oct.–Dec. 2006.

[11] Elgammal, D. Harwood, and L. Davis, "Non-parametric model for background subtraction," in Proc. IEEE Frame-Rate Workshop, 2000.

[12] Elgammal, R. Duraiswami, D. Harwood, and L. S. Davis, "Background and foreground modeling using non-parametric kernel density estimation for visual surveillance," Proc. IEEE, vol. 90, no. 7, pp. 1151–1163, Jul. 2002.

[13] S. Chen, J. Zhang, Y. Li, and J. Zhang, "A hierarchical model incorporating segmented regions and pixel descriptors for video background subtraction," IEEE Trans. Ind. Inf., vol. 8, no. 1, pp. 118–127, Feb. 2012.

[14] Haritaoglu, D. Harwood, and L. S. Davis, "W4: Real-time surveillance of people and their activities," IEEE Trans. Pattern Anal. Mach. Intell., vol. 22, no. 8, pp. 809–830, Aug. 2000.

[15] T. Zhao and R. Nevatia, "Bayesian human segmentation in crowded situations," in Proc. CVPR, 2003.

[16] L. Dong, V. Parameswaran, V. Ramesh, and I. Zoghlami, "Fast crowd segmentation using shape indexing," in Proc. ICCV, 2007.

[17] M. Kafai and B. Bhanu, "Dynamic Bayesian networks for vehicle classification in video," IEEE Trans. Ind. Inf., vol. 8, no. 1, pp. 100–109, Feb. 2012.

[18] R. Cutler and L. Davis, "Robust real-time periodic motion detection, analysis, and applications," IEEE Trans. Pattern Anal. Mach. Intell., vol. 22, no. 8, pp. 781–796, Aug. 2000.

[19] L. Zhang, S. Li, X. Yuan, and S. Xiang, "Real-time object classification in video surveillance based on appearance learning," in Proc. IEEE Int. Workshop Visual Surveillance in Conjunction With CVPR, 2007.

[20] Blum and T. Mitchell, "Combining labeled and unlabeled data with cotraining," in Proc. 11th Ann