International Journal of Recent Trends in Engineering & Research (IJRTER)
Conference on Electronics, Information and Communication Systems (CELICS'17)
Special Issue; March - 2017 [ISSN: 2455-1457]
DOI : 10.23883/IJRTER.CONF.20170331.039.35SCM

ANOMALY DETECTION USING MARKOV-MODULATED POISSON PROCESS FOR VIDEO SURVEILLANCE

S.Giriprasad[1] M.Reshma[2], A.Rifaya Banu[3], V.G.Suganthi [4],Santhosh Kumar .S[5] ,
[1] Assistant Professor, [2-5] Student,Department of Electronics and Communication Engineering,
Department of Electronics and Communication Engineering,
Coimbatore Institute of Engineering and Technology, Tamilnadu, India

ABSTRACT

Video anomaly detection plays a critical role for intelligent video surveillance. We present on abnormal video event detection system that considers both spatial and temporal contexts. To characterize the video, we first perform the spatio-temporal video segmentation and then propose a new region based descriptor called "Motion Context", to describe both motion and appearance information of the spatio-temporal segment. For anomaly measurements, we formulate the abnormal event detection as a matching problem, which is more robust than statistic model based methods, especially when the training data set is of limited size. For each testing spatio-temporal segment, we search for its best match in the training data set, and determine how normal it is using a dynamic threshold. To speed up the search process, compact random projections are also adopted. Experiments on the benchmark data set and comparisons with the state-of-the-art methods validate the advantages of our algorithm.

The data set of densely crowded pedestrian walkways is introduced and used to evaluate the proposed anomaly detector. Experiments on this and other data set show that the latter achieves state-of-the-art anomaly detection results.

I. INTRODUCTION

Nowadays, a large number of surveillance cameras have been installed due to the decreasing costs of video cameras. Intelligent video surveillance [1] is of great interests in industry applications due to the increasing demand to reduce the manpower of analysing the large-scale video data. Key technologies have been developed for intelligent surveillance, such as object tracking [2],[3], pedestrian detection [4] gait analysis [5], vehicle template recognition [6], privacy protection [7], face and iris recognition [8], video summarization [9], and crowd counting [10]. In this paper, we focus on video anomaly detection (also named as outlier detection), i.e. detecting the irregular patterns that are different from the regular video events in a given data set [11],[20] and we intend to build an abnormal event detection system that can work in crowded scenes as well. Despite many previous work of detecting video anomalies [11],[20] few of them can work well in crowded scenes, due to the following challenges:

First, a crowded scene usually contains a large number of moving persons; thus can easily distract the local anomaly detector. It is difficult, even for human beings, to effectively identify all abnormal behaviours in real time. Second, whether an event is normal or abnormal usually application and context dependent, thus it is difficult to model the abnormal event. An event may be considered as normal in one scenario while abnormal in another scenario. For real applications, it is desired that we can adaptively define the video anomaly rather than manually do this for each scenario. Third, although it is easy to obtain training videos of normal video events, it is difficult to collect sufficient samples of abnormal video events. Such an unbalanced

Special Issue; March - 2017 [ISSN: 2455-1457]
DOI:

training data brings challenges to build a robust video anomaly detector.

We illustrate two spatio-temporal video anomalies in Fig. 1 as the majority of vehicles follow the green trajectories, the U-turn moving is treated as abnormal. In Fig. 2, each ellipse stands for a moving pedestrian, but the behaviour of the red one is different from its neighbourhood, thus is considered as an abnormal event. To make a video anomaly detection system easy to use, the detection added new subsection to compare the influence of different image patch size using of video anomaly should be adaptive to different scenes.

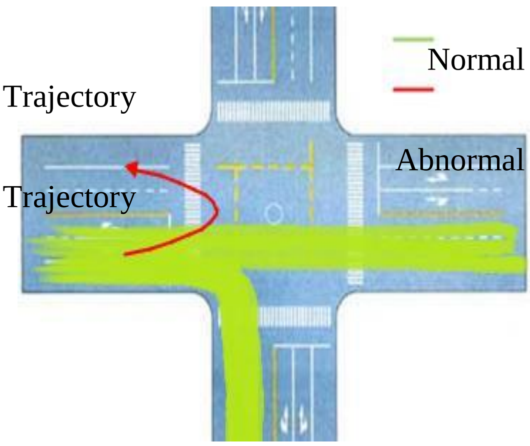

Fig.1 Temporal Abnormal Event

.

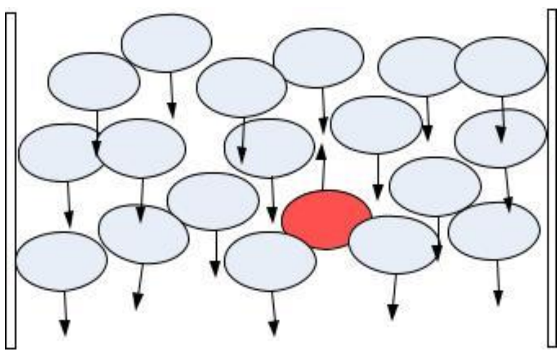

Fig.2 Spatial Abnormal Event

Fig.1. Examples of video anomalies in different scenarios.

Surveillance video is extremely tedious to monitor when events that require follow-up have very low probability. For crowded scenes, this difficulty is compounded by the complexity of normal crowd behaviours. This has motivated a surge of interest in anomaly detection in computer vision. However, this effort is hampered by general difficulties of the anomaly detection problem. One fundamental limitation is the lack of a universal definition of anomaly. For crowds, it is also infeasible to enumerate the set of anomalies that are possible in a given surveillance scenario. This is compounded by the sparseness, rarity, and discontinuity of anomalous events, which limit the number of examples available to train an anomaly detection system.

One common solution to these problems is to define anomalies as events of low probability with respect to a probabilistic model of normal behaviour. This enables a statistical treatment of anomaly detection, which conforms with the intuition of anomalies as events that deviate from the expected. However, it introduces a number of challenges. First, itomalies dependent on the makes an scale at which normalcy is defined. A normal behaviour at a fine visual scale may be perceived as highly anomalous when a larger scale is considered, or vice versa. Hence, normalcy models must be defined at multiple scales. Second, different tasks may require different models of normalcy. For instance, a detector of freeway speed limit violations will rely on normalcy models based on speed features. On the other hand, appearance is more important for the detection of carpool lane violators, i.e., single-passenger vehicles in carpool lanes. Third, crowded scenes require normalcy models robust to complex scene dynamics, involving many independently moving objects that occlude each other in complex ways, and can have low resolution.

Special Issue; March - 2017 [ISSN: 2455-1457]
DOI:

II.RELATED WORKS

Anomaly Detection:

We start by proposing an anomaly detector that accounts for scene appearance and dynamics, spatial and temporal context, and multiple spatial scales.

Mathematical Formulation:

A classical formulation of anomaly detection, which we adopt in this work, equates anomalies to outliers. A statistical model p_Xð**x**Þ is postulated for the distribution of a measurement **X** under normal conditions. Abnormalities are defined as measurements whose probability is below a threshold under this model. This is equivalent to a statistical test of hypotheses:

H_0: **x** is drawn from p_Xð**x**Þ;

H_1: **x** is drawn from an uninformative distribution p_Xð**x**Þ / 1. The minimum probability of error rule for this test is to reject the null hypothesis H_0 if p_Xð**x**Þ < , where is the normalization constant of the uninformative distribution. As usual in the literature, we consider the problem of anomaly detection from localized video measurements **x**, where **x** is a spatiotemporal patch of small dimensions.

Spatial versus Temporal Anomalies:

The normalcy model p_Xð**x**Þ can have both a temporal and a spatial component. Temporal normalcy reflects the intuition that normal events are recurrent over time, i.e., previous observations establish a contextual reference for normalcy judgments. Consider a highway lane where cars move with a certain orientation and speed. Bicycles or cars heading in the opposite direction are easily identified as abnormal because they give rise to observations **x** substantially different from those collected in the past. In this sense, temporal normalcy detection is similar to background subtraction [21]. A model of normal behaviour is learned over time, and measurements that it cannot explain are denoted temporal anomalies.

Spatial normalcy reflects the intuition that some events that would not be abnormal per se are abnormal within a crowd. Since the crowd places physical or psychological constraints on individual behaviour, behaviours feasible in isolation can have low probability in a crowd context. For example, while there is nothing abnormal about an ambulance that rides at 50 mph in a stretch of highway, the same observation within a highly congested highway is abnormal. Note that the only indication of abnormality is the difference between the crowd and the object at the time of the observation, not that the ambulance moves at 50 mph. Since the detection of such abnormalities is mostly based on spatial context, they are denoted spatial anomalies. Their detection does not depend on memory. Instead, it is based on a continuously evolving, instantaneously adaptive, definition of normalcy. In this sense, the detection of spatial anomalies can be equated to saliency detection [22].

Roles of Crowds and Scale

Most available background subtraction and saliency detection solutions are not applicable to crowded scenes, where backgrounds can be highly dynamic. In this case, it is not sufficient to detect variations of image intensity, or even optical flow, to detect anomalous events. Instead, normalcy models must rely on sophisticated joint representations of appearance and dynamics. In fact, even such models can be ineffective. Since crowds frequently contain distinct sub-entities, for example,vehicles or groups of people moving in different directions, anomaly detection requires modeling multiple video components of different appearance and dynamics. A model that has been shown successful in this context is the mixture of DTs [23]. This is the representation adopted in this work.

Another challenging aspect of anomaly detection within crowds is scale. Spatial anomalies are usually detected at the scale of the smallest scene entities, typically people. However, a normal event at this scale may be

Special Issue; March - 2017 [ISSN: 2455-1457]
DOI:

anomalous at a larger scale, and vice versa. For example, while a child that rides a bicycle appears normal within a group of bicycle riding children, the group is itself anomalous in a crowded pedestrian sidewalk. Local anomaly detectors, with small regions of interest, cannot detect such anomalies. To address this, we represent crowded scenes with a hierarchy of MDTs that cover successively larger regions. This is done with a computationally efficient hierarchical model, where MDT layers are estimated recursively.

A similar challenge holds for temporal anomalies. While their detection is usually based on a small number of video frames, certain anomalies can only be detected over long time spans. For example, while it is normal for two pedestrian trajectories to converge or diverge at any point in time, a cyclical convergence and divergence is probably abnormal. Anomaly detection across time scales is, how-ever, more complex than across spatial scales, due to constraints of instantaneous detection and implementation complexity. Since video has to be buffered before anomalies can be detected, large temporal windows imply long detection delays and storage of many video frames. Due to this, we do not consider multiple temporal scales in this work. A single scale is chosen, using acceptable values of delay and storage complexity, and used throughout our experiments. Note that, like their spatial counterparts, temporal anomaly maps are computed at multiple spatial scales. Hence, in what follows, the term "scale" refers to the spatial support of anomaly detection, for both spatial and temporal anomalies.

III. PROPOSED SCHEME

MARKOV-MODULATED POISSON PROCESS

A Markov modulated Poisson Process (MMPP) is a Poisson process whose rate varies according to a Markov process. The non-homogeneous MMPP developed in this article is a natural model for point processes whose events combine irregular bursts of activity with predictable (e.g. daily and hourly) pat-terns. We show how the MMPP may be viewed as a superposition of unobserved Poisson processes that are activated and deactivated by an unobserved Markov process. The MMPP is a continuous time model which may also be viewed as a discretely indexed non-stationary hidden Markov model by viewing intervals between events as a sequence of dependent random variables. The HMM representation allows one to probabilistically reconstruct the latent Markov and Poisson processes using a set of forward-backward recursions. The recursions allow MMPP parameters to be estimated either by an EM algorithm or by a rapidly mixing Markov chain Monte Carlo algorithm which uses the recursions for data augmentation [23]. The Markov-Poisson cascade (MPC) is an MMPP whose underlying Markov process obeys certain restrictions which uniquely order the event rates for the observed process. The ordering avoids a possible label switching issue without slowing down the rapidly mixing algorithms we use to implement the model. We apply the MPC to a data set containing click rate data for individual computer users browsing through the World Wide Web. Because the complete data posterior distribution for the MPC is a product of exponential family distributions we are able to incorporate data from multiple users into a hierarchical model using existing methods from hierarchical Poisson regression.

The Markov modulated Poisson process (MMPP) is a doubly stochastic Poisson process whose rate varies according to a Markov process. This article decomposes the MMPP into a superposition of latent Poisson processes which are activated and deactivated by a latent Markov process. The result is a natural model for point process data where events combine predictable patterns with irregular bursts of

Special Issue; March - 2017 [ISSN: 2455-1457]
DOI:

activity. The MMPP is most frequently seen in queuing theory (Du 1995; Olivier and Walrand 1994) but it has other interesting applications. Davison and Ramesh (1996) applied a discretized MMPP to a binary time series of precipitation data by numerically optimizing the discretized MMPP likelihood.

Scott (1998) used the MMPP to model criminal intrusions on a telephone network. Other uses of the MMPP exist in environmental, medical, industrial, and sociological research. Inference for MMPP parameters has received little attention because most applications of the MMPP assume known model parameters. Turin (1996) proposed an EM algorithm for finding maximum likelihood estimates of MMPP parameters [23]. Scott(1999) provided a Bayesian method for inferring the parameters of a stationary two state MMPP. This article extends Scott (1999) to the nonstationary case with an arbitrary number of states. To our knowledge this is the first treatment of inference for the non-homogeneous MMPP. We show how the MMPP can be viewed as a superposition of latent Poisson processes, which in turn may be expressed as a non-homogeneous, discretely indexed hidden Markov model (HMM) by partitioning time into intervals between observed events. Expressing the MMPP as an HMM allows one to probabilistically reconstruct the latent Markov and Poisson processes using a set of forward-backward recursions. The recursions allow MMPP parameters to be estimated through familiar latent variable methods such as the EM algorithm or MCMC data augmentation.

The Markov-Poisson cascade (MPC) is a special case of MMPP that enforces an ordering of the state space of the underlying Markov process. The MPC maintains the ordering in a natural way, so that a potential label switching issue is avoided without slowing down our rapidly mixing MCMC algorithms or modifying the specification of model parameters as in Robert and Titterington (1998).

We assume user i in the click rate dataset follows an MPC with parameters $\varphi_i = \{\lambda_{im}(t), \beta_{im}(t), \delta_{im}(t) : m = 0, \ldots, M-1\}$. For convenience we set $M = 3$, where the three states indicate respectively the absence of a Web session, a session with a slow click rate, and rapid clicking. In a serious application we would allow M to depend on i using any of several Bayesian methods for model selection or model averaging. We force $\beta_0(t) = \delta_0(t) = 0$ so that N_0 remains active as a baseline to catch isolated events. Each user's click stream almost certainly contains strong daily and hourly patterns, but these patterns are inestimable because only a single day has been observed. Therefore we assume all rates are constant, e.g. $\lambda_{im}(t) = \lambda_{im}$. Had a much longer time window been observed we could incorporate daily and hourly patterns into the click rates using the timing model proposed by Lambert *et al.* (2001), for example The familiar exponential family distributions underlying the MPC make it easy to embed MPC parameters in a hierarchical model. We assume the prior distribution

$p(\varphi_i, \ldots, \varphi_n) = \mathrm{Ga}(\lambda_{im} \mid a_{\lambda m}, b_{\lambda m})\mathrm{Ga}(\beta_{im} \mid a_{\beta m}, b_{\beta m})\mathrm{Ga}(\delta_{im} \mid a_{\delta m}, b_{\delta m})$

where $\mathrm{Ga}(\cdot \mid a, b)$ is the gamma distribution with mean a/b and variance a/b^2. The hyper parameters in are interpretable as prior event counts and observation times. For example $a_{\beta m}$ is a prior number of births for N_m, and $b_{\beta m}$ is a prior amount of time spent waiting for N_m to be born. Following Christiansen and Morris (1997), we assume an improper uniform prior on each a/b and assume each $p(a) = z_0/(z_0 + a)^2$, which is a proper normalized distribution with no moments. Christiansen and Morris show this prior has good frequency properties in the context of Poisson regression. The only "tuning parameter" is z_0, which we set to the relatively uninformative value of 0.10. Let z_i denote the missing indicators required to compute L_{com} for user i and let α denote the set of $(a_{..}, b_{..})$ pairs in (4). We used an MCMC algorithm which cycles between sampling from $p(z_i \mid \varphi_i)$, $p(\varphi_i \mid z_i, \alpha)$ for each i, and from $p(\alpha|\varphi_1, \ldots, \varphi_n)$ with $n = 1025$. We ran the algorithm for 5000 iterations and

Special Issue; March - 2017 [ISSN: 2455-1457]
DOI:

removed the first 1000 as burn-in. Figure 3 shows the remaining 4000 draws of φ_i for a sample account. The time series plots and autocorrelation functions in Figure 3 indicate rapid mixing attributable to the forward-backward recursions used in the data augmentation step.

IV. CONCLUSION

In this work, we proposed an anomaly detector that spans time, space, and spatial scale, using a joint representation of video appearance and dynamics and globally consistent inference. For this, we modeled crowded scenes with a hierarchy of MDT models, equated temporal anomalies to background subtraction, spatial anomalies to discriminant saliency, and integrated anomaly scores across time, space, and scale with a CRF. It was shown that the MDT representation substantially outperforms classical optical flow descriptors, that spatial and temporal anomaly detection are complementary processes, that there is a benefit to defining anomalies with respect to various normalcy contexts, i.e., in anomaly scale space, and that it is important to guarantee globally consistent inference across space, time and scale.

We have also introduced a challenging anomaly detection data set, composed of complex scenes of pedestrian crowds, involving stochastic motion, complex occlusions, and object interactions. This data set provides both frame-level and pixel-level ground truth, and a protocol for the evaluation of anomaly detection algorithms. The proposed anomaly detector was shown effective on both this and a number of previous data sets. When compared to previous methods, it outperformed various state-of-the-art approaches, either in absolute performance or in terms of the tradeoff between anomaly detection accuracy and complexity.

REFERENCES

1. V. Singh and M. Kankanhalli, "Adversary aware surveillance systems," Information Forensics and Security, IEEE Transactions on, vol. 4, no. 3,PP552–563, 2009.
2. S. Avidan, "Ensemble tracking," IEEE transactions on pattern analysis and machine intelligence, pp. 261–271, 2007.
3. Z. Khan and I. Gu, "Joint feature correspondences and appearance similarity for robust visual object tracking," Information Forensics and Security, IEEE Transactions on, vol. 5, no. 3, pp. 591–606, 2010.
4. N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in CVPR, 2005, pp. 886–893.
5. L. Wang, "Abnormal walking gait analysis using silhouette-masked flow histograms," in ICPR, vol. 3, 2006, pp. 473–476.
6. S. Wang and H. Lee, "A cascade framework for a real-time statistical plate recognition system," Information Forensics and Security, IEEE Transactions on, vol. 2, no. 2, pp. 267–282, 2007.
7. X. Yu, K. Chinomi, T. Koshimizu, N. Nitta, Y. Ito, and N. Babaguchi, "Privacy protecting visual processing for secure video surveillance," in ICIP, 2008, pp. 1672–1675.
8. U. Park and A. Jain, "Face matching and retrieval using soft biometrics," Information Forensics and Security, IEEE Transactions on, vol. 5, no. 3, PP406–415, 2010.
9. Y. Cong, J. Yuan, and J. Luo, "Towards scalable summarization of consumer videos via sparse dictionary selection," Multimedia, IEEE Transactions on, vol. 14, no. 1, pp. 66–75, 2012.
10. Y. Cong, H. Gong, S. Zhu, and Y. Tang, "Flow mosaicking: Real-time pedestrian counting without scene-specific learning," in CVPR, 2009, PP1093–1100.

Special Issue; March - 2017 [ISSN: 2455-1457]
DOI:

11. H. Zhong, J. Shi, and M. Visontai, "Detecting unusual activity in video," in CVPR, 2004.
12. Cheriyadat A.M and R. J. Radke, "Detecting Dominant Motions in Dense Crowds," Selected Topics In Signal Processing, IEEE Journal Of, vol. 2, no. 4, pp. 568–581, 2008.
13. Loy.A, T. Xiang, and S. Gong, "Detecting and discriminating behavioural anomalies," Pattern Recognition, vol. 44, no. 1, pp. 117–132, 2011.
14. Y. Benezeth, P.Jodoin, V. Saligrama, and C. Rosenberger, "Abnormal events detection based on spatio-temporal co-occurences," in CVPR, 2009.
15. E. S. I. R. D. Adam, A.; Rivlin, "Robust real-time unusual event detection using multiple fixed-location monitors," TPAMI, vol. 30(3) Volume 30, pp. 555 – 560, 2008.
16. V. Mahadevan, W. Li, V. Bhalodia, and N. Vasconcelos, "Anomaly detection in crowded scenes," in CVPR, 2010.
17. M. S. Ramin Mehran, Alexis Oyama, "Abnormal crowd behaviour detection using social force model," in CVPR, 2009.
18. O. Boiman and M. Irani, "Detecting irregularities in images and in video," in ICCV, 2005.
19. Y. Cong, J. Yuan, and J. Liu, "Sparse Reconstruction Cost for Abnormal Event Detection," in IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), 2011, pp. 3449–3456.
20. V. Saligrama and Z. Chen, "Video anomaly detection based on local statistical aggregates," in CVPR. IEEE, 2012, pp. 2112–2119.
21. C. Stauffer and W. Grimson, "Adaptive Background Mixture Models for Real-Time Tracking," Proc. IEEE Conf. Computer Vision and Pattern Recognition, 1999.
22. L. Itti, C. Koch, and E. Niebur, "A Model of Saliency-Based Visual Attention for Rapid Scene Analysis," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 20,
23. S. Giriprasad, S. Mohan, S.Gokul, " Video Anomaly Localization and Detection using Markov- Modulated Poisson Process, International Journal of Cntrol Theory and Applications, PP 1527-1538,2015.