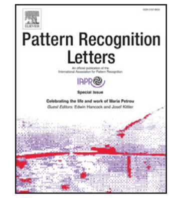
Pattern Recognition Letters 000 (2017) 1–7

Contents lists available at ScienceDirect

Pattern Recognition Letters

journal homepage: www.elsevier.com/locate/patrec



Anomaly detection based on maximum a posteriori

Shifeng Li [a,*], Chunxiao Liu [a], Yuqiang Yang [a]

Department of Electrical and Computer Engineering, Bohai University, Jinzhou, China

A R T I C L E I N F O

Article history:
Available online xxx

MSC:
41A05
41A10
65D05
65D17

Keywords:
Anomaly detection
MAP
Grid template,

A B S T R A C T

In this paper, we propose a novel method to detect abnormal events from videos based on a maximum a posteriori (MAP). Conventional methods consider the events with low-probability with respect to a model of normal behavior as anomaly. Different from the traditional approaches, the anomaly detection is achieved by a MAP estimation in our framework. The prior knowledge is obtained from the background subtraction due to the fact that the anomalies often occur at the locations consisting of moving objects, and the likelihood function is computed by comparing the similarity between the testing samples and a designed maximum grid template. Experiments on three public databases show that our method can effectively detect abnormal events in complex scenes.

1. Introduction

With increasing of videos in the world, video surveillance systems are being increasingly used in many applications. It is necessary to develop effective algorithms to aid in the automatic analysis of video data.

The abnormal events can be identified as irregular events from normal ones. Conventional methods [13,14,19] detect testing samples with lower probability as anomalies by fitting a probability model over the training data. Recently, sparse coding based scheme is applied to anomaly detection [2,11,20,28] and shows great potential. However, it is strange that the proposed approaches do not consider the prior knowledge in their framework, which can improve the anomaly detection accuracy. In this paper, we argue that the prior knowledge is important and useful for anomaly detection. For example, it is impossible that one person walks on the trees, the plane can not fly on the street, the car can not move on the water, and so on. With the prior knowledge, we can easily distinguish the anomalies in some cases. In this paper, we propose a novel framework based on maximum a posteriori (MAP), which can integrate the prior knowledge into the anomaly detection. In our work, the anomaly detection is estimated as a MAP problem, where the prior distribution is obtained from background subtraction due to the fact that the anomalies have nothing with the background and they all occur among the objects. The likelihood function is obtained by a designed maximum grid template, which can effectively capture the location, statistic information of the training samples.

The rest of this paper is organized as follows. Section 2 provides a brief overview of previous works on anomaly detection. The detailed explanation of our method is provided in Section 3. Section 4 demonstrates the effectiveness of the proposed algorithm in the published datasets, followed by the conclusions in Section 5.

2. Related works

Many methods are proposed in abnormal events detection. The first kind of approaches are based on trajectories [4,20] where the object tracking technology is employed in anomaly detection. For example, Basharat et al. first use object tracking information to construct the scene model using gaussian mixture model (GMM), and then adopt the constructed model to detect abnormal motion patterns which conflict with the trends observed in the training data. Mo et al. [20] develop a joint sparsity model and introduce the kernel technology into this sparsity model for anomaly detection. Due to the complex scenes, object tracking is not reliable in densely crowed scenes and it is far likely to lead to unsatisfactory anomaly detection results.

Owing to the limitation of object tracking based methods, many other methods are proposed to using low-level features for anomaly detection. These methods can be further divided into model based methods and rule based methods.

The model based approaches try to build one normal model based on low-level feature of the normal training samples and the behaviors of low compatibility with the model will be considered as the anomalies. The Markov Random Field (MRF) and Hid-

* Corresponding author.
E-mail addresses: lishifeng2007@gmail.com (S. Li), yangyuqiang@bhu.edu.cn (Y. Yang).

http://dx.doi.org/10.1016/j.patrec.2017.09.001

den Markov Model (HMM) models are always used to characterize the distribution of normal motion patterns. Kim and Grauman [13] propose a space-time MRF model to detect anomalies based on a grid local region graph. Benezeth et al. [5] estimate a co-occurrence matrix within the spatio-temporal volumes and use it as a potential function in a MRF model to describe the probability of observations. Kratz and Nishino [14] construct a motion-pattern distribution to capture the variations of local spatio-temporal motion patterns and use a distribution-based HMM to describe the motion transitions, and improve the framework by constructing a coupled HMM. Zhang et al. [30] propose a semi-supervised HMM framework, where the normal event models are learned from a large amount of training data, while the anomaly event models are learned by Bayesian adaptation. Ouivirach et al. [21] first build an initial set of models to explain the behaviors occurring in a small bootstrap dataset, and then use the new sequence to incrementally update the sufficient statistics of the HMM it is assigned to. Mehran et al. [19] use a Social Force Model (SFM) to estimate the interaction forces based on the moving particles obtained with the spatio-temporal feature of optical flow, and classify frames as normal and abnormal by using a bag-of-words model.

Rule based methods usually establish one rule for the normal events, if a testing sample broke the rule and it would be classified into an anomaly. Sparse based methods [3,11,16,20,29,33] can be considered as one kind widely used rule based approach. The sparse-based approaches can detect abnormal events effectively under the assumption that the normal events can be constructed from the normal basis. For example, Zhao et al. [33] propose a fully unsupervised dynamic sparse coding approach for detecting unusual events in videos based on online sparse reconstruction. Cong et al. [11] introduce the sparse reconstruction cost (SRC) over a normal dictionary to measure the normalness of the testing sample using a multi-scale histogram of optical flow (MHOF) feature. Xu et al. [29] apply the sparse coding algorithm to the dynamic texture for anomaly detection. Lu et al. [16] propose one fast method using sparse combination at a speed of 140–150 frames per second to detect anomalies. Clustering scheme is often employed to limit the amount of memory and computation. Roshtkhari and Levine [22] first adopt an online fuzzy clustering approach to group the spatio-temporal volumes into several clusters, and then calculate the ensembles of volumes to capture the contexture information with a probability framework. Finally, the anomaly detection is achieved based on the clustering of the ensembles. Saligrama and Chen [25] propose a probabilistic framework to account for local spatio-temporal anomalies. Zhang et al. [32] employ the locality sensitive hashing filters (LSHF) for anomaly detection, and achieve an encouraging performance.

Deep learning method is also employed in anomaly detection [18,23,24]. For example, Sakurada and Yairi [24] demonstrate that the auto encoders can not only be useful as nonlinear techniques, but also detect subtle anomalies. Sabokrou et al. [23] employed a sparse auto encoder to capture the local and global descriptors for the video properties. Medel and Savakis [18] develop a generative model based on a composite Convolutional Long Short-Term Memory (Conv-LSTM) network for anomaly detection. A review modelling representation of video feature for anomaly detection is proposed in [10]. Deep learning methods are popular in video-based tasks, owing to its ability to produce good representations with raw input. But this methods requires a high configuration of computer hardware and the training process needs much more training data and time for one representation.

Some of the methods mentioned above, such as model based methods, often employ the local structure information to improve the anomaly detection accuracy. However, most methods seldom take the prior knowledge into consideration directly. In this paper, we try to use the prior knowledge to guide the anomaly detection, which is integrated into a MAP framework for a more accurate detection. The experiments on public datasets show that the performance is really improved by the prior knowledge.

3. Our work

To integrate the prior knowledge into anomaly detection, we address the anomaly detection within a Baysian framework, in which the goal is to determine a posterior probability $p(\theta^t|z^t)$,

$$p(\theta^t|z^t) = \frac{p(z^t|\theta^t)p(\theta^t)}{\int p(z^t|\theta^{t'})p(\theta^{t'})d(\theta^{t'})}, \tag{1}$$

where θ^t is the anomaly state for the tth frame, and z^t is the observed low level feature at the frame t. Let $\theta^t = [x, y, p]$, where (x, y) denotes the location at which the anomalies happen, and p is the probability of the anomaly occurring at this location. $p(\theta^t)$ is the prior knowledge for the anomaly, $p(z^t|\theta^t)$ is the likelihood function measuring how likely the anomaly happens at the tth frame. The optimal state $\hat{\theta}^t$ can be obtained by MAP over the current frame features:

$$\hat{\theta}^t = \arg\max_{\theta^t} p(z^t|\theta^t)p(\theta^t). \tag{2}$$

In this paper, the anomaly detection is based on cubes, that is, our goal is to estimate the probability of each cube using MAP. Furthermore, for similarity, we do not consider the former frames before the tth frame. Instead, we use a ‘box’ filter to smooth the anomaly detection results. The top flowchart of our algorithm is shown in Fig. 1.

3.1. Prior knowledge

The $p(\theta^t)$ in Eq. (1) stands for the prior knowledge for the tth frame in anomaly detection. In fact, the person can not walk on the trees, and the trees can not move, the anomaly always happens when the object moves very quickly, and so on. To indicate the prior knowledge, we use the background subtraction to compute $p(\theta^t)$, as we think that the anomalies only happen with the moving objects not with the background. There are many methods for the background subtraction, such as MOG background subtraction [7], subspace learning background subtraction [6], Robust Principal Component Analysis (RPCA) background subtraction [8] and so on. In this paper, we use the RPCA for the background subtraction. The algorithm used in our experiments is the Fast Principle Component Pursuit (FPCP) due to its effectiveness and efficiency.

RPCA method decomposes the frames into low-rank L and sparse S components. The background is modeled by low-dimensional subspace L and moving objects belong to the S component. Suppose the sparse approximation is S^t for the tth frame in the testing video. We compute the binary map as the final $p(\theta^t)$:

$$p(\theta^t) = S^t > \varepsilon, \tag{3}$$

where ε is a hard threshold obtained by $\varepsilon = 0.5 * std(S)^2$, and $std(\cdot)$ is the standard deviation function. It is obvious that if an anomaly happens, it will occur at location containing the moving objects. Fortunately, the binary prior map of S^t can provide the accurate location for the anomaly detection.

3.2. Likelihood function using maximum grid template

The likelihood function $p(z^t|\theta^t)$ is a function of the parameter θ^t, and describes the probability we see the observation z^t provided the state θ^t at the tth frame, which is important for a MAP problem. To obtain the likelihood function $p(z^t|\theta^t)$, we design a maximum grid template to measure the similarity between the testing samples and the template. The training samples provided