Pattern Recognition 51 (2016) 443–452

Contents lists available at ScienceDirect

Pattern Recognition

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

Combining motion and appearance cues for anomaly detection

Ying Zhang [a], Huchuan Lu [a,*], Lihe Zhang [a], Xiang Ruan [b]

[a] School of Information and Communication Engineering, Dalian University of Technology, Dalian 116023, China
[b] OMRON Corporation, Kusatsu, Japan

ARTICLE INFO

Article history:
Received 14 April 2015
Received in revised form
14 August 2015
Accepted 7 September 2015
Available online 28 September 2015

Keywords:
Anomaly detection
Motion model
Appearance model
Support Vector Data Description (SVDD)

ABSTRACT

In this paper, we present a novel anomaly detection framework which integrates motion and appearance cues to detect abnormal objects and behaviors in video. For motion anomaly detection, we employ statistical histograms to model the normal motion distributions and propose a notion of "cut-bin" in histograms to distinguish unusual motions. For appearance anomaly detection, we develop a novel scheme based on Support Vector Data Description (SVDD), which obtains a spherically shaped boundary around the normal objects to exclude abnormal objects. The two complementary cues are finally combined to achieve more comprehensive detection results. Experimental results show that the proposed approach can effectively locate abnormal objects in multiple public video scenarios, achieving comparable performance to other state-of-the-art anomaly detection techniques.

1. Introduction

Intelligent video surveillance [1,2] has been attracting more and more attention in recent years. Among the many research objectives, anomaly detection from video sequences plays an important role in discovering various irregularities, such as restricted-area access [3], wrong direction or route [4–6], people carrying cases [7], falling [8–10], group fighting and panics [11], a car making an illegal U-turn [12,13], jaywalkers [12], and some other unusual events.

Despite the fact that anomaly detection has been successfully used in intelligent transportation systems, security alarm systems for houses, offices and public places, it is still encountering a series of challenges, which can be summarized as follows: (1) it is hard to define representative normal regions in observed video streams; (2) the boundary between normal and abnormal objects is often ambiguous; (3) the exact notion of anomaly is rather subjective and changes over different applications; (4) it is tricky to label the abnormal behaviors for videos; (5) normal behaviors tend to keep evolving; (6) observed data may contain noise.

A great diversity of approaches have been proposed to solve one or more problems mentioned above. Most of them were designed to work for specific scenarios, where different representations of motion and appearance were analyzed with different models. A large majority of these methods detect abnormalities following two main assumptions:

- Abnormal events rarely occur in video sequences [14–17].
- Abnormal events have low similarities to normal events [4,18–20].

In this paper we consider properties of abnormal instances from the perspective of human cognition. Unusual objects are rare things with unexpected appearance or motion patterns. In order to generically detect anomalies, we argue that any abnormal object has at least one of the following characteristics:

- Motion anomaly, like objects moving with an unexpected speed or in a wrong direction, along unplanned routes, at undesired locations, etc.
- Appearance anomaly, such as strange postures or unidentified objects.

Many abnormal objects have both the above characteristics. In this paper, we propose a novel framework by exploring the motion and appearance cues for anomaly detection. For appearance anomaly detection, we employ spatio-temporal gradients as appearance descriptors and develop a novel scheme based on Support Vector Data Description (SVDD) [21], which obtains a spherically shaped boundary in the non-linearly transformed feature space to include most of the normal samples, while test examples falling out of the boundary would be considered as unusual. For motion anomaly detection, we utilize the optical flow vectors to characterize motion features and describe the distribution of normal motion patterns via histogram statistics, where we introduce the notion of "cut-bin" analogous to the boundary in SVDD. A new observation would be regarded as anomalous if it crosses the line of "cut-bin". The overall abnormal objects are located according to

* Corresponding author. Tel./fax: +86 411 84708971.
E-mail address: lhchuan@dlut.edu.cn (H. Lu).

http://dx.doi.org/10.1016/j.patcog.2015.09.005
0031-3203/

the results of both motion and appearance anomaly detection. The proposed method is robust for discovering multiple unusual events in different scenarios. Fig. 1 shows the overview of our approach.

2. Related work and paper contents

2.1. Related work

Many approaches have been proposed for abnormality detection. According to whether or not samples of normal and abnormal activities are needed for training or initialization before anomaly detection, they can be broadly classified into three fundamental types: supervised, semi-supervised and unsupervised approaches.

The first type of the anomaly detection methods requires the labels available for both normal and abnormal samples, which is generally involved in training classifiers. These methods are customarily designed for specific abnormal behaviors whose properties are predefined by people, such as falling detection [9,22,10], fighting detection [23] and traffic violation detection [12,13]. Velocity and trajectories act as two most widely used cues to classify normal and abnormal cases.

The second type of approaches only needs normal data for training, which is the most popular technique adopted by researchers. These methods can be further divided into two subcategories, rule based methods and model based methods.

Rule based approaches attempt to establish a rule that any sample which breaks the rule would be considered as irregular. Sparse coding [24,4,25] is widely used in rule based approaches, with a larger reconstruction cost indicating a higher probability of being abnormal. Online dictionary updating [4] has been added into the coding framework to handle the concept drift, while Lu et al. [25] learned sparse combinations to speed up the coding process, reaching a speed of 140–150 frames per second on average. However, the detection results are greatly affected by the threshold, which often varies over different scenes. Similarities based methods are also exploited by many researchers [26,18–20], abnormal score of a test sample is measured by its similarity to the training samples. Clustering [20] and sub-class discovering [19] were adopted to speed up the computation. While the differences between normal activities are usually not obvious, clustering might not work as well as expected.

Model based approaches try to build a model for normal behaviors, while instances of low probability with respect to the model will be rejected as anomalies. The most widely used models are Markov Random Field (MRF) model [14,16] and Hidden Markov Model (HMM) [27–30], which are extended in various applications. Kim et al. [14] introduced a space–time MRF model characterizing distribution of normal motion patterns, while Benezeth et al. [16] used MRF to describe the co-occurrence distribution of normal observations. Kratz et al. [27] modeled the temporal and spatial relationship between local spatio-temporal motion patterns using distribution-based HMM and coupled HMM respectively. Andrade et al. [29] grouped the video segments into several different classes using spectral clustering and trained a Multiple Observation Hidden Markov Model (MOHMM) for each class. Some other models are also designed for anomaly detection, such as social force model (SFM) proposed by Mehran et al. [5] for describing individual motion dynamics and interaction forces in crowds, and mixture of dynamic texture (MDT) [31] modeling normal occurrences both in space and time. Adam et al. [15] presented a simple and fast algorithm for unusual events' detection based on optical flow histogram statistics of normal activities of multiple fixed-location monitors. While Wu et al. [6] learned a normal scene model using chaotic invariants of Lagrangian particle trajectories to characterize crowded scene. The energy model was also used in predicting irregularities. Cui et al. [11] proposed an interaction energy potential function to model the action and interaction with the surroundings of normal objects changing over time, and events with sudden changes in functions are likely to be unusual. Entropy calculated on the spatiotemporal information of the interest points is used to measure disorders in [32]. Kwon et al. [33] developed a graph editing framework with a predefined energy model whose parameters reflect causality, frequency, and significance of events, which can be used for event summarization and rare event detection. The energy based methods are often sensitive to multiple parameters, and a particular energy function can only function well for specific scenes.

The third type of approaches require neither normal nor abnormal examples in advance and the anomaly detection is based

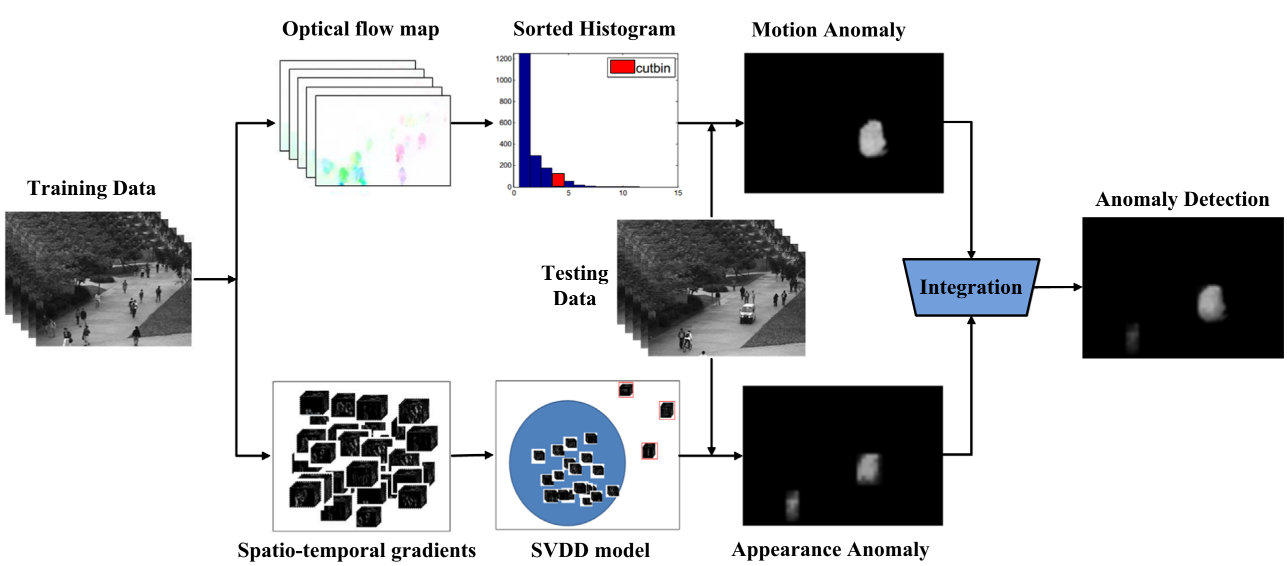

Fig. 1. Overview of anomaly detection algorithm. For motion anomaly detection, optical flow vectors are computed and aggregated into sorted histograms, cut-bin of the histogram for each location is calculated to judge abnormalities. For appearance anomaly detection, we extract spatio-temporal gradients from normal samples and train a SVDD model to identify anomalies.