Signal Processing: Image Communication 60 (2018) 6–12

Contents lists available at ScienceDirect

Signal Processing: Image Communication

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

Anomaly detection based on two global grid motion templates

Shifeng Li *, Yuqiang Yang, Chunxiao Liu

Department of Electronic Engineering, Bohai University, JinZhou, China

ARTICLE INFO

Keywords:
Anomaly detection
Grid motion template

ABSTRACT

In this paper, we propose a novel method to detect abnormal events from videos based on two global grid motion templates (GGMTs) which are able to capture the motion distribution, space and scale information. The GGMTs contain the maximum and minimum grid motion templates which can effectively distinguish the anomalies from the normal motion distribution. One GGMT is composed of several non-overlap local grid motion templates with each one corresponding to a special location. Each local grid motion template is represented by a motion histogram obtained by computing the maximum/minimum motion distribution from the training samples. Experiments on the public datasets show that our method can effectively detect abnormal events in complex scenes.

1. Introduction

The abnormal events can be identified as irregular events from normal ones. Conventional methods [1–3] detect testing samples with lower probability as anomalies by fitting a probability model over the training data. Recently sparse coding scheme is applied to the anomaly detection [4–7] and shows great potential. The mentioned methods usually construct one model or dictionary according to the low-level features, and cannot accurately capture the motion distribution for all locations. In fact, the motion distribution should not be the same at different locations or scales. In stead of constructing one model or dictionary for anomaly detection, we capture the motion distribution for each grid location and scale using the histogram technique. In this paper, we show that we can achieve the state-of-the-art performance via employing the motion feature based on our proposed two GGMTs, containing the maximum GGMT and minimum GGMT. For one scale GGMT, the motion histogram for each grid volume location is first computed and then the maximum histogram is constructed as the local grid template at this location. All local grid templates constitute the GGMT, which contains not only the motion distribution, but also the space and scale information. The maximum GGMT can capture the upper bound motion patterns of the training samples, while the minimum GGMT can characterize the lower bound normal motion distribution. The anomaly detection is carried out by computing the similarity between templates and testing samples. The advantages of our method are as follows:

- Different from the traditional method which only use one model or dictionary of all locations for anomaly detection, we propose a scheme that constructs every local grid motion template at each location for a precise motion distribution representation.
- We propose two motion templates, including maximum grid motion template and minimum grid motion template for anomaly detection. Besides, our method has a very fast speed for training and testing due to the simple max and min operations.

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 conclusions in Section 5.

2. Related works

Many methods are proposed in anomaly detection. Based on the feature used in anomaly detection, the methods can be divided into several categories. The first type of approaches is based on trajectories [8,7] using object tracking technology. For example, Basharat et al. [8] first used object tracking information to construct a scene model using Gaussian Mixture Model (GMM), and then used the constructed model to detect abnormal motion patterns which conflicted with the trends observed in the training data. Mo et al. [7] developed a joint sparsity model and introduced kernel technology into this sparsity model for anomaly detection. Due to the complex scenes, object tracking is not reliable

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

http://dx.doi.org/10.1016/j.image.2017.09.002
Received 25 March 2017; Received in revised form 5 September 2017; Accepted 5 September 2017
Available online 8 September 2017
0923-5965/

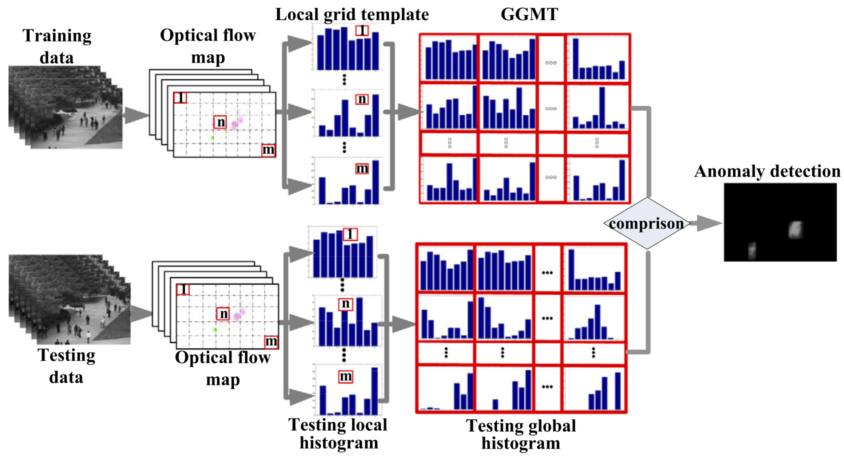

Fig. 1. The top flowchart of our method.

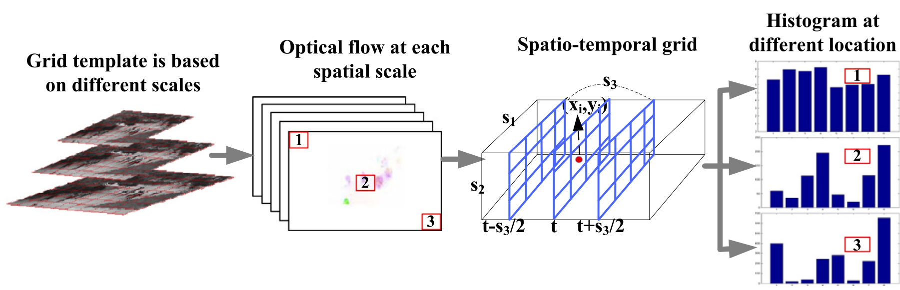

Fig. 2. Examples for grid template. The grid template is computed at different scales and extracted in the spatio-temporal grid.

in densely crowed scenes and it is far likely to lead to unsatisfactory anomaly detection.

Owing to the limitation of object tracking based method, the second category of algorithms [9,2,10,11] focuses on low-level motion features, such as optical flow [9,1,12], histogram-based feature [2,10,11], and so on. The optical flow patterns were detected in frames and the maximum a posteriori was estimated to identify the abnormal events using the learned Markov Random Field (MRF) [1]. Optical flow feature in each pair of consecutive frames was extracted for motion histogram estimation in each segmented region for abnormal events detection in [12]. Wang et al. proposed a novel unsupervised learning framework to model activities and interactions in crowded and complicated scenes based on the optical flow feature. Cong et al. [4] proposed to detect abnormal events via a sparse reconstruction over the normal bases using Multi-scale Histogram of Optical Flow (MHOF). Zhao et al. [13] proposed a fully unsupervised dynamic sparse coding approach to detect unusual events in videos based on an atomically learned event dictionary, using Histogram of Oriented Gradient (HOG) and Histogram of Optical Flow (HOF) extracted on the Spatio-temporal Interest Points (STIPs). Bertini et al. [10] proposed an approach for anomaly detection and localization based on spatio-temporal features that capture scene dynamic statistics together with appearance using the histogram scheme. Javan et al. [14] presented a novel approach for video parsing and online learning of anomalous behaviors in surveillance videos, based on densely sampled spatio-temporal video volumes characterized by the histogram in polar coordinates. Lu et al. [11] used a simple gradient feature for anomaly detection at a speed of 140–150 frames per second on average.

The third category is based on blob-based feature. Since this method considers the foreground pixel blobs as events, it always uses the background subtraction to remove background influence. For example, Xiang and Gong [15] first used an adaptive Gaussian mixture background model to detect foreground pixel and then grouped the detected foreground pixels into blobs to represent the basic events. Li et al. [16] also performed background subtraction and detected image events as foreground pixel blobs.

The other category employs the texture-based feature. This approach extracts the texture from the video frames to capture the normal texture patterns for abnormal events detection. For example, Mahadevan et al. [17] proposed a model of normal crowd behavior based on mixtures of dynamic textures and labeled the outliers under this model as anomalies. Li et al. [18] also used the mixture of dynamic textures for temporal and spatial anomaly detection. Xu et al. [19] presented a novel approach for unusual events detection via sparse reconstruction of dynamic textures described by Local Binary Patterns (LBP) from three orthogonal planes. The LBP and sparse coding were also used in [20] for anomaly detection in nature images.

3. Our work

The top algorithm flowchart for one scale is shown in Fig. 1. In this paper, the training video is first resized into different scales, and then divided into spatio-temporal grid of size $s_1 \times s_2 \times s_3$ around each location at each spatial scale (see Fig. 2). Finally we compute the corresponding grid template in each grid to comprise the GGMT at each scale, which can capture space, scale and motion distribution of the training data. The anomaly detection probability is obtained by comparing the similarity between the trained GGMT and the testing video.

3.1. Low-level motion feature

As shown in Section 2, the low-level motion feature is important in anomaly detection due to its effectiveness and practicability. Therefore, we also focus on the anomaly detection based on low-level motion feature. In our work, we use a very simple but effective motion feature for anomaly detection. We first use the method [21] to compute the optical flow map, and then represent each spatio-temporal volume on