Neurocomputing ■ (■■■■) ■■■–■■■

Contents lists available at ScienceDirect

Neurocomputing

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

# Summarizing surveillance videos with local-patch-learning-based abnormality detection, blob sequence optimization, and type-based synopsis

Weiyao Lin [a,*], Yihao Zhang [a], Jiwen Lu [b], Bing Zhou [c], Jinjun Wang [d], Yu Zhou [e]

[a] Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai, China
[b] Advanced Digital Sciences Center, Singapore, Singapore
[c] School of Information Engineering, Zhengzhou University, Zhongyuan, China
[d] Institute of Artificial Intelligence and Robotics, Xi'an Jiao Tong University, Xi'an, China
[e] Institute of Information Engineering, Chinese Academy of Sciences, Beijing, China

ARTICLE INFO

*Article history:*
Received 4 October 2014
Received in revised form
7 December 2014
Accepted 17 December 2014
Communicated by Tao Mei

*Keywords:*
Video synopsis
Blob sequence optimization
Abnormality detection

ABSTRACT

In this paper, we propose a new approach to detect abnormal activities in surveillance videos and create suitable summary videos accordingly. The proposed approach first introduces a patch-based method to automatically model normal activity patterns and key regions in a scene. In this way, abnormal activities can be effectively detected and classified from the modeled normal patterns and key regions. Then, a blob sequence optimization process is proposed which integrates spatial, temporal, size, and motion correlation among objects to extract suitable foreground blob sequences for abnormal objects. With this process, blob extraction errors due to occlusion or background interference can be effectively avoided. Finally, we also propose an abnormality-type-based method which creates short-period summary videos from long-period input surveillance videos by properly arranging abnormal blob sequences according to their activity types. Experimental results show that our proposed approach can effectively create satisfying summary videos from input surveillance videos.



## 1. Introduction

Video surveillance is of increasing importance in many applications including traffic control, unusual alarming [1–6]. In many scenarios, people need to browse videos to find events of interest or perform analysis. However, since surveillance videos are usually long, it is laborious to watch the entire videos. Thus, it is essential to create short-period summary (or abstract/synopsis) videos which summarize important events in long-period surveillance videos. In this way, people's labor can be greatly saved by only watching these short summary videos [1,2,14,38]. Therefore, in this paper, we focus on creating suitable summary videos for input surveillance videos.

First, since most people are interested in abnormal activities in surveillance videos, detecting abnormalities in videos is crucial in analyzing and summarizing surveillance videos. Many algorithms have been proposed on abnormality detection [3,5,6,17–19,27–36]. However, most of these works only focus on detecting abnormalities while the differentiation of abnormality types is seldom addressed. In practice, differentiating abnormal activity types is important in creating well-organized summary videos.

Second, it is also important to extract accurate foreground blob[1] sequences for objects such that objects can be suitably separated and arranged to create satisfying summary videos. Although many tracking algorithms have been proposed [7–10,12,16,37], their performances are still less satisfactory due to the interferences from object occlusion or complex background. Besides, most tracking-based methods only focus on achieving object bounding boxes while the suitable segmentation of object blobs is not addressed. In practice, achieving accurate object blob is non-trival in creating satisfying summary videos.

Third, creating suitable summary videos from long surveillance videos is another key issue. Recently, video synopsis methods [1,2,24] were proposed which extracted and put together object

* Corresponding author. Tel.: +86 21 34208843; fax: +86 21 34204155.
*E-mail address:* hellomikelin@gmail.com (W. Lin).

[1] In this paper, a blob refers to a connected foreground region for one or several objects [10], as in Fig. 4(c).

http://dx.doi.org/10.1016/j.neucom.2014.12.044
0925-2312/

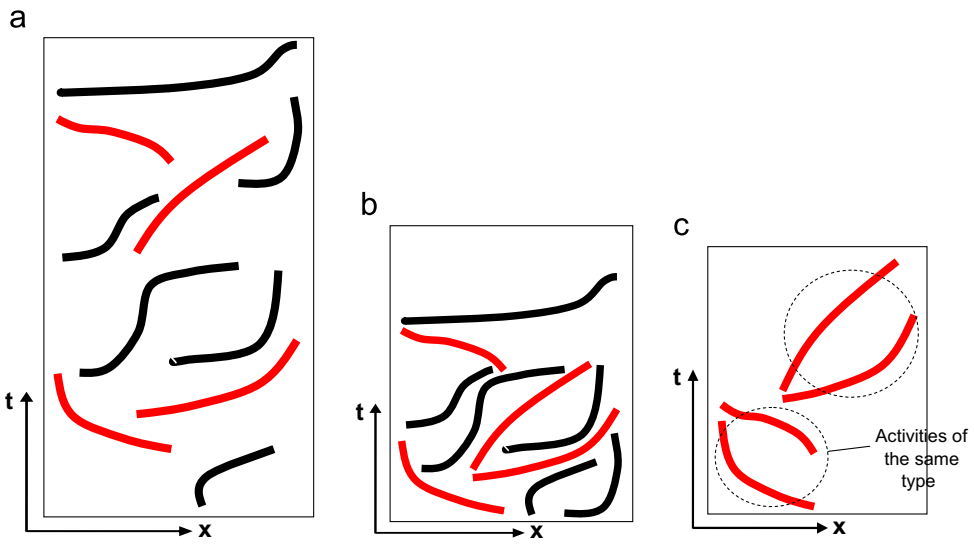


**Fig. 1.** (a): An input video including object trajectories (i.e., blob sequences); (b): the synopsis video by [1] which moves and puts together object trajectories from different periods; (c): the synopsis video by our approach which only performs synopsis on abnormal trajectories and put together trajectories of the same type, i.e., put together trajectories which start from the same region and end in the same regions into the same time period. (*Note*: *t* represents the time domain and *x* represents the spatial domain of a video. The red lines represent abnormal trajectories and the black lines represent normal trajectories, best viewed in color). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

blob sequences from different periods, as in Fig. 1(b). Video synopsis has the advantage of creating short summary videos while suitably maintaining all blob sequences of interest. However, the existing synopsis methods still have the following limitations: (1) They are still less effective in summarizing crowd-scene videos where the huge number of blob sequences will make the synopsis videos chaotic and less understandable. (2) They only focus on compressing the length of videos while seldom consider the proper arrangement of similar activity types.

In this paper, we propose a new approach to detect abnormal activities from surveillance videos and create suitable summary videos accordingly. The contributions our approach can be summarized as follows:

(1) We introduce a patch-based method to automatically model normal activity patterns and use them to detect abnormal activities. Besides, based on the observation that each scene should include "key regions" and all activity trajectories in a scene should go through part of them (as in Fig. 2), we also propose to extract key regions from a scene and use them to classify abnormal activities into different types (i.e., activities are classified into the same type when they pass through the same key regions, as in Fig. 2). By introducing key regions, we are not only able to improve the abnormality detection accuracy, but are also able to organize abnormalities into different types which enables the creation of well-organized summary videos in later steps.
(2) Based on the assumption that most people are interested in abnormal activities in surveillance videos, we propose an abnormality-type-based video synopsis method which summarizes surveillance videos by only synopsizing over abnormal blob sequences. Moreover, the proposed method further introduces an activity-type cost during the synopsis process such that blob sequences of the same activity type (i.e., activities passing through the same key regions) can be arranged closely in summary videos. With this method, we are able to create well-organized summary videos even for crowded scenes, as in Fig. 1(c).
(3) We also propose a blob sequence optimization process which integrates spatial, temporal, size, and motion correlation among objects to extract suitable foreground blob sequences for abnormal objects. With this process, blob extraction errors due to occlusion or background interferences can be effectively avoided.

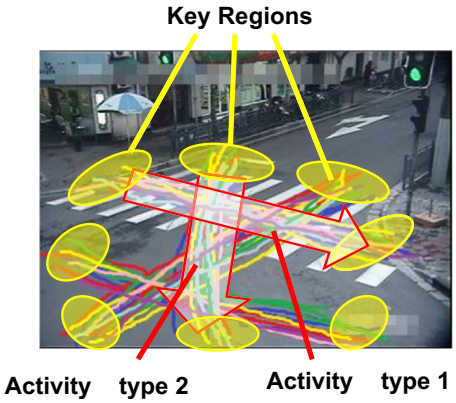


**Fig. 2.** An example of key regions (yellow circles) and activity types (red arrows). Key regions refer to regions in a scene where all activity trajectories in the scene should go through part of them (e.g., cluster of trajectory terminals), and activity types refer to activity classes whose trajectories pass through the same key regions (best viewed in color). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

The rest of the paper is organized as follows: Section 2 discusses the related work. Section 3 describes the framework of our proposed approach. Sections 4–6 describe the details of our approach. Section 7 shows the experimental results. Section 8 concludes the paper.

## 2. Related works

Since abnormal activity detection is one of the most important issues in surveillance video analysis, it has attracted a lot of research works [3,5,6,17–19,27–36]. Many people detected abnormalities by parsing the motion trajectories of objects. For example, Zelniker

Please cite this article as: W. Lin, et al., Summarizing surveillance videos with local-patch-learning-based abnormality detection, blob sequence optimization, and type-based synopsis, Neurocomputing (2015), http://dx.doi.org/10.1016/j.neucom.2014.12.044