Poster: Crowdsourcing for Video Traffic Surveillance

Hui Wen[1], Qiang Li[1], Qi Han[2], Shiming Ge[1],Limin Sun[1]*
[1]Beijing Key Laboratory of IOT Information Security Technology, Institute of Information Engineering, CAS, China
[2]Department of EECS, Colorado School of Mines, Golden, CO USA 80401
E-mails:{wenhui,liqiang43,geshiming,sunlimin}@iie.ac.cn, qhan@mines.edu

Categories and Subject Descriptors

H.4.m [**Information Systems Applications**]: Miscellaneous

Keywords

Crowdsourcing, Video traffic surveillance, Event detection

1. INTRODUCTION

Video traffic surveillance monitors traffic situations such as traffic jams, traffic accidents, or running a red light. Although automatic traffic event detection has been studied for years, current systems often fail to handle various situations and do not fully take advantage of existing video traffic surveillance data. Hence, there is a need for an approach that integrates labor resources with intelligent video analysis to enhance the robustness of video analysis models and fulfill the demands of traffic surveillance.

Motivated by the intuition that a driver or pedestrian often needs to know the exact traffic conditions before selecting a particular route, we propose a crowdsourcing [2] surveillance framework to assist existing traffic surveillance systems. In particular, people can use their smartphones to check the detected traffic situation and the corresponding video clips received from the video surveillance system, and make quick judgements about the received results. This fine-grained information provided by traffic surveillance system not only shows the detected traffic results but also presents live video clips. Furthermore, smartphone users can provide their feedback to the system for improving the intelligent video surveillance model or correcting errors that may be present in the current traffic event detection.

2. FRAMEWORK

Figure 1 shows our proposed framework. Basically, the system informs smartphone users of abnormal traffic situations and gets the feedback from users to improve the traffic event detection model. This crowdsourcing allows the surveillance tasks to be distributed among a large number of smartphone users who provide ground truth of the traffic conditions.

There are several key components in the framework. (1) *Traffic event detection.* Traffic video analysis is used to train

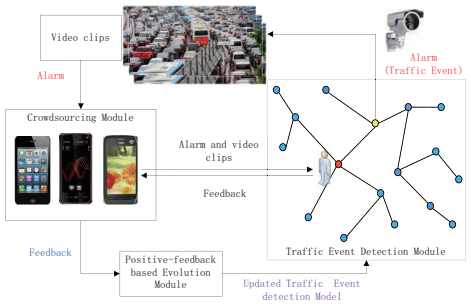

Figure 1. Overview of Proposed Framework

a model for detecting various traffic events. Our current work focuses on traffic accidents detection to validate the framework. This module applies subspace clustering or dictionary learning [1] techniques to first identify normal patterns and then detect traffic accidents based on deviation from these normal patterns. (2) *Crowdsourcing.* Different from a traditional video surveillance system, our framework allows smartphone users to get the video clips from the system via a push or pull based method. Smartphone users then become "traffic surveillance supervisors." Once they view live surveillance videos, they can provide feedback to the system to improve the robustness of the video analysis. (3) *Positive-feedback based evolution.* Simple yet valuable feedback is provided by smartphones users to indicate whether the traffic alert is right or wrong. The feedback module then updates the abnormal traffic event detection model online by positive judgment obtained from smartphone users. Specifically, the PN-learning method[3] is used for model self-learning. The method evaluates the crowdsourcing feedback with the features extracted from the video clips, identifies the motion feature that have been classified as being contradictory to structural constraints, and augments the sample set with the representative samples.

There are several key issues in implementing this framework. They include designing proper feedback strategies, developing an online abnormal traffic event detection model, solving the optimization problem of the subspace clustering or dictionary learning, considering the structure constraints of the positive and negative samples in PN-learning, and synchronization of feedback from a large number of users.

3. REFERENCES

[1] Y. Cong, J. Yuan, and J. Liu. Abnormal event detection in crowded scenes using sparse representation. *Pattern Recognition*, 46:1851–1864, 2013.

[2] J. Deng, J. Krause, and L. Fei-Fei. Fine-grained crowdsourcing for fine-grained recognition. In *CVPR*, pages 580–587, 2013.

[3] Z. Kalal, J. Matas, and K. Mikolajczyk. Pn learning: Bootstrapping binary classifiers by structural constraints. In *CVPR*, pages 49–56, 2010.

*Limin Sun is the corresponding author.

MobiSys'14, June 16–19, 2014, Bretton Woods, New Hampshire, USA.
ACM 978-1-4503-2793-0/14/06.