Expert Systems With Applications 42 (2015) 9449–9467

Contents lists available at ScienceDirect

Expert Systems With Applications

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

Automatic surveillance in transportation hubs: No longer just about catching the bad guy

Simon Denman[a,b,*], Tristan Kleinschmidt[a], David Ryan[a,b], Paul Barnes[a], Sridha Sridharan[a,b], Clinton Fookes[a,b]

[a] *Airports of the Future Project, Queensland University of Technology (QUT), GPO Box 2434, 2 George Street, Brisbane, QLD 4001, Australia*
[b] *Image and Video Laboratory, Queensland University of Technology (QUT), GPO Box 2434, 2 George Street, Brisbane, QLD 4001, Australia*

ARTICLE INFO

Keywords:
Transportation
Security
Operational performance
Video surveillance
Operational analytics
Airport operations

ABSTRACT

As critical infrastructure such as transportation hubs continue to grow in complexity, greater importance is placed on monitoring these facilities to ensure their secure and efficient operation. In order to achieve these goals, technology continues to evolve in response to the needs of various infrastructure. To date, however, the focus of technology for surveillance has been primarily concerned with security, and little attention has been placed on assisting operations and monitoring performance in real-time. Consequently, solutions have emerged to provide real-time measurements of queues and crowding in spaces, but have been installed as system add-ons (rather than making better use of existing infrastructure), resulting in expensive infrastructure outlay for the owner/operator, and an overload of surveillance systems which in itself creates further complexity. Given many critical infrastructure already have camera networks installed, it is much more desirable to better utilise these networks to address operational monitoring as well as security needs.

Recently, a growing number of approaches have been proposed to monitor operational aspects such as pedestrian throughput, crowd size and dwell times. In this paper, we explore how these techniques relate to and complement the more commonly seen security analytics, and demonstrate the value that can be added by operational analytics by demonstrating their performance on airport surveillance data. We explore how multiple analytics and systems can be combined to better leverage the large amount of data that is available, and we discuss the applicability and resulting benefits of the proposed framework for the ongoing operation of airports and airport networks.

1. Introduction

The scale of challenges facing society in providing more advanced critical infrastructure, and in particular transportation hubs, is substantial. An increased importance has been placed (and will continue to be placed) on transportation hubs to accommodate increased demand as cities continue to expand, and global air travel becomes more accessible. Consequently, the complexity of these hubs is increasing; airports are a perfect example where a multitude of factors are continually in play, including technological advancements, changes in regulations, and the interaction of multiple stakeholders including (but not limited to) government agencies, airport operators, airlines, security contractors, commercial operators, and of course, the travelling public (Ashford, Coutu, & Beasley, 2013). With this increase in demand and complexity, it can be argued that it is becoming more important to monitor the operational performance of the system. This is especially true for many privatised airports who often depend on non-aviation revenue sources to a greater extent than traditional aviation revenue.

Unfortunately, with millions of passengers passing through these hubs on a daily basis, the infrastructure itself become prime targets for terrorist activities (Tsai, Rathi, Keikintveld, Ordonez, & Tambe, 2009). Recent examples of such attacks on transportation hubs include the London train bombings in 2005, the Glasgow Airport car bombing in 2007, the Domodedovo International Airport bombing in 2011; the Peshawar Airport attack in 2012 and the Jinnah International Airport attack in 2014. With this constant threat in place, it is extremely important to ensure transportation hubs are safe and secure for all involved.

[*] Corresponding author at: Airports of the Future Project, Queensland University of Technology (QUT), GPO Box 2434, 2 George Street, Brisbane, QLD 4001, Australia. Tel.: +61 731389329.

E-mail addresses: s.denman@qut.edu.au (S. Denman), kleinschmidt_tf@yahoo.com.au (T. Kleinschmidt), david.ryan1@gmail.com (D. Ryan), p.barnes@qut.edu.au (P. Barnes), s.sridharan@qut.edu.au (S. Sridharan), c.fookes@qut.edu.au (C. Fookes).

http://dx.doi.org/10.1016/j.eswa.2015.08.001
0957-4174/© 2015 Elsevier Ltd. All rights reserved.

Over the past decade, surveillance cameras have become commonplace in public locations, including transportation hubs (Adrem, Dell'orto, & Lennerman, 2007; Wells, Allard, & Wilson, 2006; Welsh & Farrington, 2009). This is a direct result of an increased focus on public safety and security, but can also be attributed to the reduced cost of cameras and their associated computing infrastructure which help to keep overall security costs down (Adrem et al., 2007). The increase in surveillance cameras has given rise to an increase in computer systems to manage, and in many cases to analyse the incoming camera feeds. At present, these feeds are typically monitored by staff, and detecting events as they happen is very challenging due to the sheer amount of data being presented to each operator.

To assist human operators in monitoring large CCTV networks, there has been a significant increase in computer vision research and development, to create algorithms to analyse and extract information from the CCTV feeds. These developments have tended to focus on security related tasks such as object/person tracking, perimeter surveillance, motion segmentation, abnormal event detection and recognition, and biometrics (e.g face, iris, fingerprint) for person identification (Fookes et al., 2010). Many of these algorithms have begun to be implemented within video management and analytic systems, giving commercial video analytic packages a wide range of (primarily) security capabilities.

Operational analytics however, such as crowd counting and queue monitoring, have received less attention and while commercial systems do exist to perform these tasks, they are fewer in number and often require specially placed cameras making them difficult to integrate with existing systems. This is despite the comparatively poor performance of security based systems, which are prone to missed detections and false alarms.

Furthermore, the integration of these emerging techniques into airports or other critical infrastructure has received limited attention. When discussing the implementation of surveillance techniques, existing research has focussed on how a single surveillance task may function within a piece of infrastructure in isolation (i.e. Li, Wu, Karanam, and Radke, 2014 considers person re-detection in an airport environment; while Arroyo, Yebes, Bergasa, Daza, and Almazán, 2015 consider suspicious behaviour detection in a shopping mall). Similarly, when considering the system wide implications of video surveillance, the literature has focussed on aspects such as the data and networking requirements of such large scale systems (Ajiboye, Birch, Chatwin, & Young, 2015; Chang, Wang, Wang, Liu, & Ho, 2012); interfaces to retrieve and display results (Ye, Liao, Dong, Zeng, & Zhong, 2015); or the needs of researchers and developers to aid in the development of such techniques (Nazare, dos Santos, Ferreira, & Robson Schwartz, 2014).

Within this paper we propose an automated surveillance framework for both operational and security tasks for on-site and across-site monitoring. Whilst ensuring that a wide range of possible surveillance technologies are included in this framework, we specifically discuss the overlap between video analytics for security and operational analytics for operational monitoring which can be exploited to make better use of CCTV networks in public spaces. We present an overview of intelligent surveillance techniques for security and operational tasks, and show that although security has long been the focus of surveillance deployments, the operational video analytics currently in development are in many ways, more appropriate for deployment.

We show, on airport surveillance data, how recent approaches can be used and combined to extract measures of operational performance such as crowd sizes, processing rates and dwell times. The performance of these approaches, as well as their strengths and weaknesses from a real-world standpoint (i.e. deployment requirements and challenges) are discussed, and we explore how these techniques can be used in tandem with other statistical modelling approaches to provide better situational awareness. To demonstrate how such a combined security and operations framework could benefit a transport hub, we develop a case study around airport security, incident response and level of service monitoring to demonstrate the potential of video analytics as a solution to both these needs.

The remainder of this paper is structured as follows: Section 2 presents an overview of intelligent surveillance and provides an outline of the current abilities of security analytics; Section 3 presents our proposed surveillance framework, and an overview of operational analytics and how they can be applied to a transport environment; Section 4 presents two case studies examining how our proposed framework could be applied to an airport environment; and Section 5 concludes the paper.

2. Automatic surveillance: a history in security

Surveillance systems are an essential and integral component in transportation networks, public places, and other critical infrastructure where it is necessary to monitor activities, threats, and to prevent or investigate criminal or other unwanted activity.

An increased focus on security coupled with falling costs of hardware has seen an increase in the number of CCTV management products available. Some systems (such as Iomniscient,[1] BlueEye Video,[2] Agility Video,[3] ObjectVideo [4]) also offer video analytics: algorithms which can extract information from the incoming feeds in real-time or near real-time. The capabilities of such products vary significantly and as such, it is helpful to categorise them through a high-level classification outlined below.

- 1st generation: Traditional analogue CCTV systems with recording facilities through tape or digital video recorders.
- 2nd generation: Highly capable "Video Management Systems" utilising large IP networks (cameras may be digital or analogue with encoders). These systems have a suite of low-level image processing tools (such as perimeter intrusion detection, loitering, abandoned object detection, etc.).
- 3rd generation: True multi-view capable intelligent surveillance systems with robust semantic information extraction.

We argue that most commercial solutions are still only 2nd generation systems (with a select few 2.5 generation) and are often characterised by high false-alarm rates, and limited knowledge of the environment in which they are deployed (i.e. camera calibration). A significant advancement in capabilities is still required before 3rd generation systems are reached, i.e. "cognitive" systems that can track, identify and explain what is taking place (Bellotto et al., 2009). This includes the development of: true multi-view capability (rather than single-view with simple camera network topologies); automatic camera calibration; robust tracking and recognition of people and events that are invariant to the challenging day-to-day operating conditions including illumination, pose, viewpoint; and invariance to noisy, cluttered complex environments. Recent advances in deep learning and convolutional neural networks indicate one direction that may advance these goals. Significant gains have been made in fields such as speech recognition (Deng & Yu, 2014), natural language processing (Manning et al., 2014), object recognition (Erhan, Szegedy, Toshev, & Anguelov, 2014) and pedestrian detection (Luo, Tian, Wang, & Tang, 2014) by leveraging very large amounts of data to automatically learn complex relationships within the data. The data requirements of deep-learning have to date meant that it's applications has been restricted to data rich domains; however it offers a promising direction to enable the development of true '3rd generation' systems.

[1] http://iomniscient.com/.

[2] http://www.blueeyevideo.com/.

[3] http://www.vidient.com/.

[4] http://www.objectvideo.com/.