Pattern Recognition 44 (2011) 117–132

Contents lists available at ScienceDirect

Pattern Recognition

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

Detecting and discriminating behavioural anomalies

Chen Change Loy *, Tao Xiang, Shaogang Gong

School of EECS, Queen Mary University of London, London E1 4NS, UK

ARTICLE INFO

Article history:
Received 20 April 2010
Received in revised form
6 July 2010
Accepted 20 July 2010

Keywords:
Anomaly detection
Dynamic Bayesian Networks
Visual surveillance
Behavior decomposition
Duration modelling

ABSTRACT

This paper aims to address the problem of anomaly detection and discrimination in complex behaviours, where anomalies are subtle and difficult to detect owing to the complex temporal dynamics and correlations among multiple objects' behaviours. Specifically, we decompose a complex behaviour pattern according to its temporal characteristics or spatial-temporal visual contexts. The decomposed behaviour is then modelled using a cascade of Dynamic Bayesian Networks (CasDBNs). In contrast to existing standalone models, the proposed behaviour decomposition and cascade modelling offers distinct advantage in simplicity for complex behaviour modelling. Importantly, the decomposition and cascade structure map naturally to the structure of complex behaviour, allowing for a more effective detection of subtle anomalies in surveillance videos. Comparative experiments using both indoor and outdoor data are carried out to demonstrate that, in addition to the novel capability of discriminating different types of anomalies, the proposed framework outperforms existing methods in detecting durational anomalies in complex behaviours and subtle anomalies that are difficult to detect when objects are viewed in isolation.

1. Introduction

The recent large-scale deployments of surveillance cameras have led to a strong demand in systems of automated anomaly detection in visual surveillance [1–3]. Earlier work is mostly focused on anomaly detection from well-defined simple behaviours in an uncrowded scenario [4,5]. More recently, the primary research focus has shifted to complex behaviour scenario in which a behaviour pattern is characterised by hierarchical temporal dynamics and/or complex correlations among multiple objects.

Anomaly detection in complex behaviours is challenging because the differences between real-life true anomalies (rather than exaggerated acts) and normal ones are often rather subtle visually and not well-defined semantically. One way to model such subtle differences is to consider that anomalies are associated with deviations in the expected temporal dynamics embedded in complex behaviours, which in turn can be considered as having layered hierarchical structures. In addition, different ways of deviations from the expected temporal dynamics lead to different types of anomalies, the discrimination of which has never been attempted to date although it is often of practical use in real-world applications. In a crowded multiple object scenario, anomaly detection becomes even more challenging because visual evidences often span across a large spatial and temporal context, anomaly is thus difficult to detect if an object is viewed in isolation.

To facilitate effective modelling and anomaly detection for complex behaviours, it is natural to decompose the modelling task into a number of sub-tasks. Most existing techniques resort to *object-based decomposition* which employs a standalone model with the model structure being factorised in accordance with the corresponding temporal processes of individual objects [6,7]. However, object-based decomposition relies on object segmentation and tracking and therefore is prone to problems associated with occlusion and trajectory discontinuities when applied to a crowded wide-area scene. In addition, object-based decomposition will lead to very complex model structure making model learning and inference intractable in the presence of large number of objects. Moreover, it offers no mechanism for discriminating different types of anomalies and reducing the effect of noise and error from the observation space.

To address these problems, we propose to perform *behaviour-based decomposition* on a complex behaviour and model the decomposed behaviours with a *cascade of Dynamic Bayesian Networks* (CasDBNs), in which a DBN model at each stage is connected to the model in the next stage via its inferential output. More specifically, behaviour-based decomposition factorises the behaviour space into sub-spaces based on directly exploring the behaviour semantics defined by different temporal characteristics of the behaviour (e.g. co-occurrence, temporal order, and temporal duration) and the spatio-temporal visual context where the behaviour occurs. Behaviours are inherently context-aware, exhibited through constraints imposed by scene layout and the

* Corresponding author. Tel.: +44 20 7882 8019; fax: +44 20 8980 6533.
E-mail addresses: ccloy@dcs.qmul.ac.uk, ccloy225@gmail.com (C.C. Loy), txiang@dcs.qmul.ac.uk (T. Xiang), sgg@dcs.qmul.ac.uk (S. Gong).

0031-3203/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.patcog.2010.07.023

temporal nature of activities in a given scene. We believe that better behaviour modelling can be achieved based on behaviour-based decomposition because the important context-awareness nature of complex behaviours is exploited explicitly, which has been largely neglected by previous object-based decomposition based approaches.

Apart from employing a different decomposition strategy, the proposed framework differs from existing approaches in that it deploys multiple DBN models in a cascade structure. This model structure is motivated by the following key observations:

(i) It is noted that different DBN models have different levels of sensitivity towards different types of anomalies. It is therefore possible to exploit this characteristic by employing a cascade of DBNs, with each of them being sensitive to one specific type of anomalies. This enables us to integrate the evidences from each DBN models to achieve a more accurate detection, and more importantly behaviour discrimination.

(ii) It is well known that noise and error in the low-level visual features are inevitable in a real-world scenario. By constructing a cascade structure with each stage being connected using the inferential output of the previous stage, the models in later stages of the cascade will be less affected by the noise and error in the observation space.

(iii) While a single model generally suffers from the scalability problem given large number of objects, a CasDBNs would benefit greatly from behaviour decomposition in avoiding this problem since the complexity of each individual model in the cascade is well controlled after the decomposition.

We present two instantiations of our framework to address two fundamental and open problems of anomaly detection in complex behaviours. In Section 4, we formulate the framework for *detecting and discriminating anomalies* by their abnormal temporal dynamics (e.g. atypical duration and irregular temporal order) embedded implicitly in the behaviour structure. In Section 5, the framework is used to address the problem of *modelling multi-object correlations* in a crowded wide-area scene and detecting subtle anomalies that are difficult to detect when objects are viewed in isolation.

1.1. Discriminating different temporal causes of anomalies

It is not only necessary but also critical to both detect and discriminate different types of anomalies based on the temporal characteristics of expected behaviours. In many real-world scenarios, there could be only one type of anomalies that are deemed as critical for triggering an alarm. For instance, in a bank branch, a different order of "entering into the branch" and "using an ATM outside the branch" is of no significance. However, the durational abnormality in front of the ATM may be of more interest. On the other hand, in a convenience store, the temporal order of "paying" and "leaving the shop" is important, whilst variations in the time spent at these atomic actions of the shopper behaviour are less critical.

In order to model and differentiate behaviours by their intrinsic characteristics, we consider a complex behaviour as a spatio-temporal pattern organised naturally in an hierarchical structure. For instance, as can be seen from Fig. 1, a person's typical behaviour in an office can include a sequence of ordered atomic actions with certain duration such as entering the office, working at a desk, printing, and leaving the office. Each atomic action itself is also composed of multiple constituents having certain duration and temporal order among them (e.g. entering the office can consist of opening the door and then walking toward the desk). A normal behaviour pattern would follow a typical order of atomic actions with certain duration. Deviation from either one or both of these temporal characteristics would cause an anomaly.

In this paper, we show that different DBN models can exhibit different levels of sensitivity given different types of anomalies. Based on this finding, we propose to decompose a complex behaviour based on different temporal characteristics, particularly the temporal order and temporal duration. This is achieved by exploiting different DBN models in a cascade, with each of

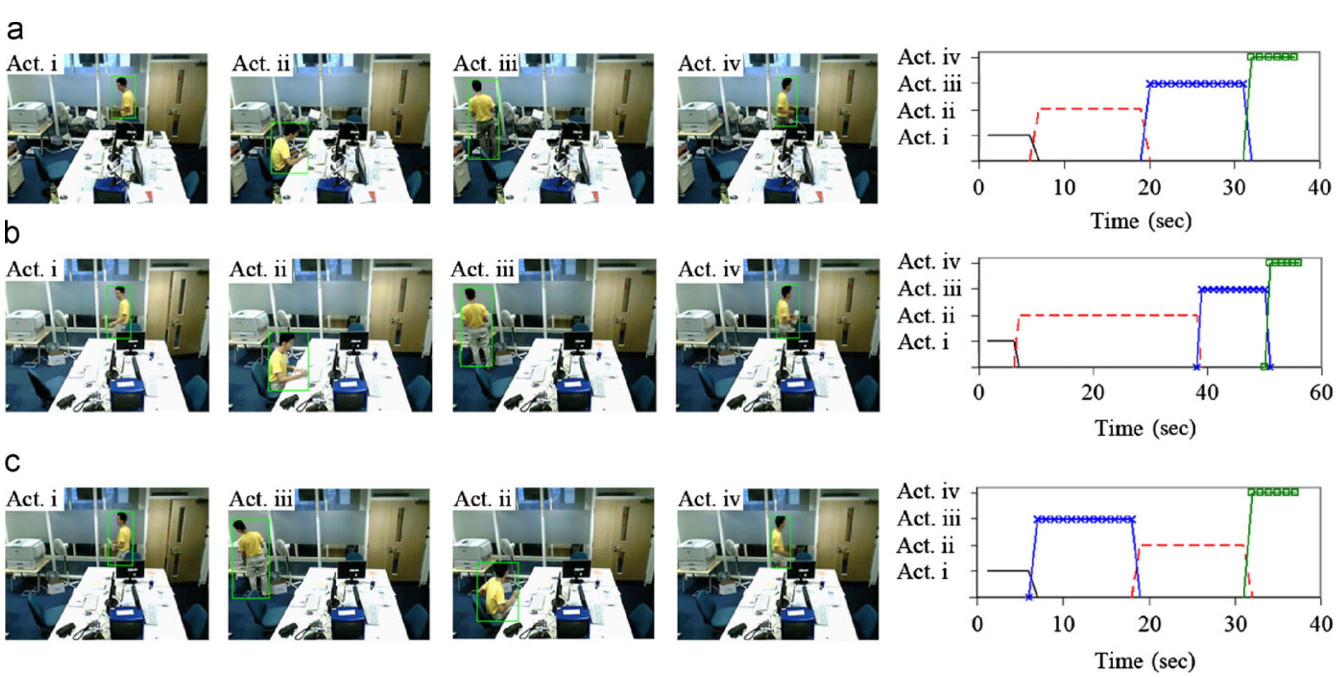

Fig. 1. Example frames of three behaviour sequences in an office environment and the associated ground truth of action occurrences. Although the behaviour sequences share the same set of atomic actions ([Act. i] entering, [Act. ii] working at a desk, [Act. iii] printing and [Act. iv] leaving), sequence (b) and sequence (c) exhibit abnormal temporal dynamics. (a) Normal behaviour sequence; (b) Behaviour sequence with atypical temporal duration; (c) Behaviour sequence with irregular temporal order.