Expert Systems with Applications 42 (2015) 7991–8005

Contents lists available at ScienceDirect

Expert Systems with Applications

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

# Expert video-surveillance system for real-time detection of suspicious behaviors in shopping malls

Roberto Arroyo *, J. Javier Yebes, Luis M. Bergasa, Iván G. Daza, Javier Almazán

*University of Alcalá, Department of Electronics, 28871 Alcalá de Henares, Madrid, Spain*

ARTICLE INFO

*Article history:*
Available online 19 June 2015

*Keywords:*
Video-surveillance in shopping malls
Background subtraction
Human tracking
Occlusion management
Appearance features
Behavioral analysis

ABSTRACT

Expert video-surveillance systems are a powerful tool applied in varied scenarios with the aim of automatizing the detection of different risk situations and helping human security officers to take appropriate decisions in order to enhance the protection of assets. In this paper, we propose a complete expert system focused on the real-time detection of potentially suspicious behaviors in shopping malls. Our video-surveillance methodology contributes several innovative proposals that compose a robust application which is able to efficiently track the trajectories of people and to discover questionable actions in a shop context. As a first step, our system applies an image segmentation to locate the foreground objects in scene. In this case, the most effective background subtraction algorithms of the state of the art are compared to find the most suitable for our expert video-surveillance application. After the segmentation stage, the detected blobs may represent full or partial people bodies, thus, we have implemented a novel blob fusion technique to group the partial blobs into the final human targets. Then, we contribute an innovative tracking algorithm which is not only based on people trajectories as the most part of state-of-the-art methods, but also on people appearance in occlusion situations. This tracking is carried out employing a new two-step method: (1) the detections-to-tracks association is solved by using Kalman filtering combined with an own-designed cost optimization for the Linear Sum Assignment Problem (LSAP); and (2) the occlusion management is based on SVM kernels to compute distances between appearance features such as GCH, LBP and HOG. The application of these three features for recognizing human appearance provides a great performance compared to other description techniques, because color, texture and gradient information are effectively combined to obtain a robust visual description of people. Finally, the resultant trajectories of people obtained in the tracking stage are processed by our expert video-surveillance system for analyzing human behaviors and identifying potential shopping mall alarm situations, as are shop entry or exit of people, suspicious behaviors such as loitering and unattended cash desk situations. With the aim of evaluating the performance of some of the main contributions of our proposal, we use the publicly available CAVIAR dataset for testing the proposed tracking method with a success near to 85% in occlusion situations. According to this performance, we corroborate in the presented results that the precision and efficiency of our tracking method is comparable and slightly superior to the most recent state-of-the-art works. Furthermore, the alarms given off by our application are evaluated on a naturalistic private dataset, where it is evidenced that our expert video-surveillance system can effectively detect suspicious behaviors with a low computational cost in a shopping mall context.



## 1. Introduction

In the current world, surveillance has become an essential element in a lot of daily activities to guarantee human security and property and assets protection. It is present in all kind of locations: banks, prisons, airports, parking lots, petrol stations, stores and any imaginable business or enterprise. Due to this, it exists an incipient need related to automating certain surveillance tasks for assisting security officers and allow them develop their work in a more efficient way.

Nowadays, the video images captured from cameras strategically located are the principal element in any surveillance system. For this reason, computer vision processing can be employed for

* Corresponding author.
*E-mail addresses:* roberto.arroyo@depeca.uah.es (R. Arroyo), javier.yebes@depeca.uah.es (J.J. Yebes), bergasa@depeca.uah.es (L.M. Bergasa), ivan.garciad@uah.es (I.G. Daza), javier.almazan@depeca.uah.es (J. Almazán).

http://dx.doi.org/10.1016/j.eswa.2015.06.016
0957-4174/

extracting useful data from these videos and reasoning this information out with the aim of automating several video-surveillance tasks by giving off alarms when risk events are detected.

This paper is focused on a specific case of video-surveillance: to detect potentially suspicious human behaviors in shopping malls. In this scenario, there are some particular situations which must be analyzed, such as store entry or exit, loitering events that can culminate in a theft or situations where a cash desk is unattended, as shown in Fig. 1.

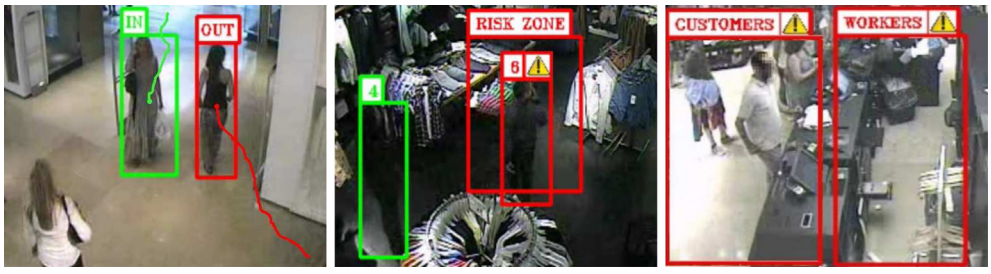
Before giving off the described alarms, there are some previous stages which must be considered in order to locate and follow the potentially suspicious people in the videos, as shown in Fig. 2. This previous process starts by making an image segmentation using background subtraction techniques. A comparison among the main background subtraction methods of the state of the art is carried out in this article with the aim of selecting the best suited for this specific application. In the next step of the process, the foreground objects, distinguished as blobs, are filtered by size and positional factors and grouped by our blob fusion algorithm. Afterwards, the tracking stage is essential for making an effective operation in the subsequent high-level phases such as trajectory analysis. Attending to this, an innovative tracking method is proposed, which consists in a two-step algorithm:

1. The detected objects to track (usually humans) are matched along the video sequence employing an optimization approach based on solving the association problem as a LSAP (Linear Sum Assignment Problem) (Easterfield, 1946) and considering the estimations of a Kalman filter (Kalman, 1960).
2. Occlusions between objects are managed applying a method based on visual appearance, in which several image descriptors are tested: GCH (Global Color Histogram) (Novak & Shafer, 1992), LBP (Local Binary Pattern) (Ojala, Pietikainen, & Harwood, 1994) and HOG (Histogram of Oriented Gradients) (Dalal & Triggs, 2005). The features extracted are compared through a metric based on SVM kernels similar to the proposed by Vapnik and Cortes (1995) and Moghaddam and Yang (2000).

The usage of the introduced expert system increases the efficacy in a video-surveillance control center for a shopping mall in comparison with traditional methods, allowing human operators manage a higher number of cameras and their corresponding risk situations. Tests executed in naturalistic conditions demonstrate the reliability of our expert video-surveillance application and its efficient performance in real-time.

The main contributions of this paper are focused on the different innovative algorithms implemented in our expert video-surveillance system. First of all, although our segmentation stage is based on previous background subtraction techniques, we have carried out a complete study to choose the most proper for our system and, besides, we have designed a new blob fusion technique that reduces typical segmentation errors. In addition, our novel tracking algorithm represents an interesting new approach with respect to the classic state-of-the-art methods which were mainly based on trajectories, because we also use visual appearance information in occlusion situations for improving the performance of our proposal compared to other related works. Finally, the alarms analyzed by our system in shopping malls are a key contribution of this paper. To the best of our knowledge, this is the first work in the state of the art that considers the specific suspicious behaviors in shops processed by our expert application. We also contribute results and comparisons that corroborate the remarkable performance of our proposal.

The rest of the paper is organized as follows: Related work is discussed in Section 2. Section 3 describes the proposed methods for video pre-processing and human tracking. The methodology implemented in our expert video-surveillance system for detecting suspicious behaviors in shopping malls and giving off the




(a) Entry and exit. (b) Loitering event. (c) Unattended cash desk.

**Fig. 1.** Alarms detected by our expert video-surveillance system in shopping malls.


(a) Background subtraction. (b) Blob fusion. (c) Tracking.

**Fig. 2.** Video pre-processing and human tracking.