Security System Based on Suspicious Behavior Detection

Enrique Bermejo, Oscar Déniz and Gloria Bueno
enric2186@gmail.com Universidad de Castilla-La Mancha, Spain.

ABSTRACT

In recent years, the demand for image analysis applications of video surveillance has grown rapidly. The latest advances in video surveillance have aimed at automating the monitoring itself, so that it is a computer (not the security personnel) what observes the images and detects suspicious behavior or events. In this context, we present system for the automatic detection of suspicious behavior in public buildings, that obtains high resolution image of the individual or individuals who have activated the alarm in the system.

1. INTRODUCTION

Most current video surveillance systems share one characteristic: they need human operator to constantly watch monitors that show the images captured by the cameras [1]. The effectiveness of these systems is not determined by their technological characteristics but by the person who is monitoring the system [2]. Today, thanks to advances in many fields of computer vision, these systems are evolving to become virtually automatic. It is not human observer who detects suspicious situations, but algorithms that process the captured images and detect suspicious behavior or events [3].

The purpose of this article is to describe semiautomatic video surveillance system that is able to detect suspicious situations using artificial vision, as well as to facilitate the operator's work by generating visual and audible alerts. Most surveillance cameras installed today have static location and capture low quality images (we all have once seen on television news the images of an assault at shop where the action and the individual appears in small area of the image). Despite the high quality of the available tools for image processing, typical captured images are not very useful for crime investigation. The increasing need for security in public spaces makes real-time detection of suspicious behavior essential, rather than simply recording them [4].

This system aims to control the camera movement and zoom in order to obtain higher quality pictures for further investigations. The system relies on two physical components: security camera to get pictures of the monitorized environment and motorized Pan-Tilt.

The motorized Pan-Tilt allows the use of *patrol mode*. This patrol mode is sequence in which the camera remains static during configured time monitorizing defined area/position. When this time is over, the system changes the position of the camera and re-starts the monitorization (of the new area). Figure 1 represents the patrol mode described.

Section 2 describes the image processing, behavior analysis and alerting methods developed. Then, Section 3 shows the results obtained with the system. Finally, Section 4 details the most important conclusions.

2. DEVELOPED SYSTEM

2.1 Image Processing

During the initial processing of the images obtained by the surveillance camera, the system performs a segmentation of objects in order to obtain a data set that provide the position and movement they make. In order to get the data set, we use back-ground subtraction techniques ([5], [6]) and motion analysis ([7]), using the OpenCV library [8].

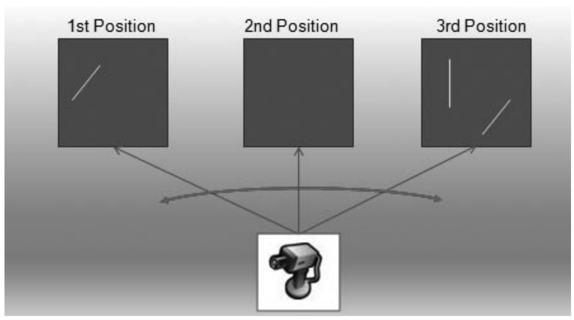

Figure 1. Patrol Mode Representation.

2.1.1 Background Subtraction

For modeling and object detection the system uses the background subtraction method known as Running Average. This method involves the creation of a background model from the average of the *n* previous frames, as described in [9]. Although this method is highly sensitive to changes in light and noise, it requires little memory and processing time, which is ideal for real-time operation.

In order to clean up the raw segmented image obtained during this process, the system realizes a connected-components analysis, in which it takes in a noisy image and then uses the morphological operations of opening and closing to eliminate noise and segment the objects that are large enough. After the analysis it is possible to nd the contours of the objects and retrieve all contours of size above a dened threshold.

2.1.2 Motion Detection

After object segmentation, the system analyzes the motion generated during the sequence captured by the camera. The technique used is Motion Templates[8], which allows recording the movement generated by the displacement of the object's dierent shapes. This generates a motion history image (MHI) of the dierent silhouettes of the object in motion.

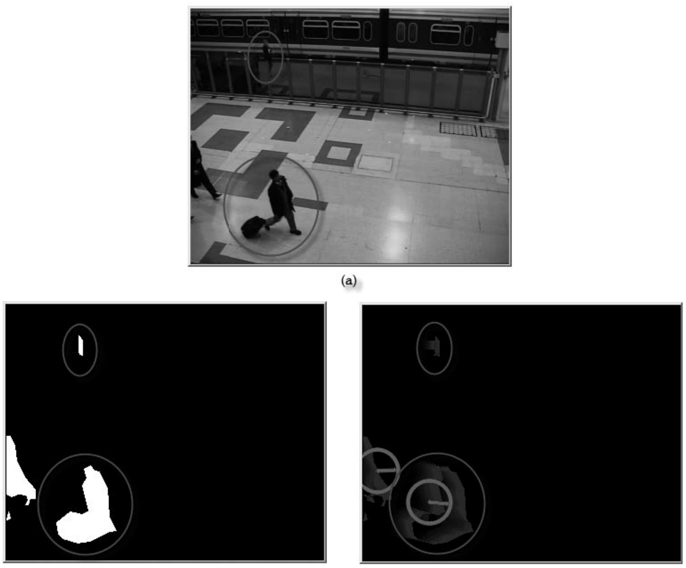

Figure 2. Motion direction (360x288). (a) Original Image. (b) Foreground. (c) MHI.

Thanks to the intensity information provided by the timestamps in the MHI, it is possible to calculate the motion gradient. That is, by analyzing the pixel values of the MHI the system is able to extract a vector indicating the flow of movement. This topic is described in detail in [10], with an analysis of the calculation of speed and orientation histograms.

This technique allows the system to obtain data that can be used for further behavioral analysis.

These data include:

1. Object's position.
2. Orientation angle.
3. Movement direction.

Figure 2 shows the results of background subtraction processes (b) and the motion template (c).

2.2 Behaviour Analysis

After obtaining the data referring to all objects that are in each frame, the system analyzes the data looking for suspicious actions such as:

- Trespassing of imaginary lines.
- Fights.
- Running people or Riots.

2.2.1 Restricted Area Detection

The trespass detection consists of the constant evaluation of the moving objects that are detected from MHI. This makes use of simple planar geometry, namely the equations of the line, the relative positions and the intersection point between two lines.

The data that allow the system to make these checks are, rst, the restricted crossing lines stored in a XML conguration le (created during system conguration), and second, the motion lines of the objects detected during the motion history processing.

The rst step is to obtain information from both the conguration le and the history of motion, then estimate the line parameters and analyze, line by line, the possibility that one of the lines detected intersects with one of the restricted crossing lines. If there is any chance of intersection, the system checks whether the trespass direction is allowed or prohibited (this is also specied in the conguration). If it detects a forbidden trespass, it calculates the exact spot of intersection.

If the intersection point is within the detection limits (start point and end point) of the restricted line drawn, the motion detected is considered suspicious and, therefore, there is a trespass risk. In this case, the system generates an alert only when the moving object is at a distance less than a dened threshold. Moreover, with the information of the object that triggered the alarm, a region of interest (ROI) is established around this object and tracking is activated.

2.2.2 Races, Fights and Riot Detection

The detection of races, ghts and riot consists of the evaluation of positions and motion angles of the moving objects. In order to do this, for each frame captured by the surveillance camera, the system calculates the change, from one frame to the next, the po sitions of every object and the change of motion angles associated with those objects. To perform these calculations we made use of the Hausdor metric.

The Hausdor metric measures the distance between two compact subsets of a metric space. This metric basically represents the maximum distance of a set of points with the closest point of another set. The formal denition of the Hausdor distance, dH (X, Y), is as follows:

$$d_H(X,Y) = max\left\{ \underset{x\in X}{sup} \, \underset{y\in Y}{inf} \; d(x,y), \underset{y\in Y}{sup} \, \underset{x\in X}{inf} \; d(x,y) \right\}$$

Where X and Y are two subsets of points in the metric space M.

This metric has many applications in computer vision and can be used to locate a particular template (pattern) in an image [11], in tracking and classifying objects, in comparison of 2D images with 3D objects. In this work we use it to obtain the degree of movement of an object in an image sequence. More specifically, the Hausdor distance is used for calculating the deviation of the positions and motion angles of the moving objects detected with the Motion Templates algorithm. The system uses a modied version of the Hausdor algorithm, where the values are normalized by calculating the average of the minimum distances obtained.

When the distance measured by this algorithm is small, it means that the positions and directions of movement components have not changed much from one frame to another. When the distance is large, the variation is abrupt and may be indicative of rapid linear movement and/or changing direction, which we considered indicative of possible run, riot or ght.

2.3 Alert Generator

In the case that suspicious activity is detected, either an trespass of prohibited areas, ght, riot or run, the system will manage a series of alarms in order to store a record of the action, and alert the operator of the system. Four types of alerts have been dened.

2.3.1 Visual Alerts

Visual alerts are divided into two groups. First we have real-time alerts, where a circle is drawn around by the individual who caused the suspicious action, as well as a line indicating its motion direction. These alerts are useful for the operator to pay attention to the area of interest in the images of the video sequence. Secondly, we have storage alerts, which are images in high resolution and with more zoom than the typical video capture. These images may be useful for further investigation.

2.3.2 Sound Alerts

The sound alerts consist of a brief sound reproduction that alerts the operator, so that he can focus on the monitor. The sound is played whenever the system detects a suspicious action.

2.3.3 Logs

Besides the generation of visual and audible alerts, when the system detects suspicious behavior it stores a record indicating the date and time, as well as the type of behavior that has been detected. This is useful because the date and type of action facilitate the reconstruction of events from the images stored as visual alerts.

2.3.4 Tracking

If the action detected is a trespass, the system enters a state in wich tracks the object that generated the alert. During this period, the system estimates the new position of the object relative to the center of the image, perform the necessary movement of the Pan-Tilt to center the object in the image, and, then, increases the camera zoom and performs a high resolution capture of the image. Tracking is done via the mechanism of Camshift Tracking (Continuously Adaptive Mean-Shift) [12]. In Figure 3 it is shown a sequence of images obtained during the tracking process.

3. EXPERIMENTAL RESULTS

The images below show some of the results obtained during the operation of the monitoring system. For those examples we have used images obtained from a FireWire surveillance camera and a video from the IEEE International Workshop on Performance Evaluation of Tracking and Surveillance (PETS), that refers to a train station monitorzation.

First, Figure 4 shows the result of a trespass (a), in which a man is crossing the imaginary line defined, the template of motion (b), the region of interest dened on the individual before start tracking on this individual (c) and, nally, the high-resolution image stored as a log for further investigation(d).

Figure 3: Tracking a moving object (512x384).

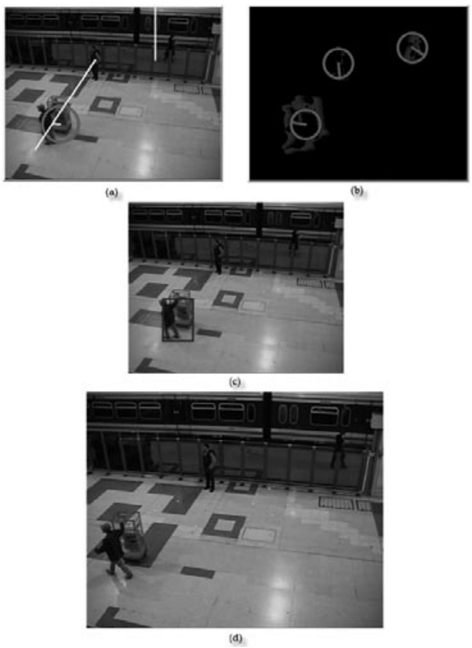

Figure 4. Trespass detected.(a),(b) y (c) (360x288). (d) (720x576).

In Figure 5 it is observed a trespass detected by the surveillance camera. Image (a) shows a snapshot of the moment in which the trespass occurs. The sequence of images (b) through (e) shows the zoom progress during tracking. After the tracking, the system also stores a high resolution image as it is shown in image (f). In this image we can observe that the individual appears at the center of the image.

Finally, Figures 6 and 7 show the results obtained using the Hausdor distance for the detection of runs and ghting, respectively. As shown in the images, the system stores images as evidence that the action has been identied as suspicious. In Figure 7 we can observe that the individuals who have activated the alarm appears at the center of the image with an increasing zoom.

Figure 5. (a) Overrun (512x384). (b), (c), (d) y (e). Tracking (512x384). (f) Detection Proof (1024x768).

4. CONCLUSION

We built a prototype of video surveillance hardware-software system to detect potentially dangerous events in real time and alert the human operator.

In addition, we have successfully fullled the goal of obtaining a view with more detail of the area or individual that has generated an alarm thanks to the pan-tilt-zoom mechanism.

The system deployment is very simple, since it enables realtime threshold setting that allow to modify the degree of detections made by the system during its execution, thus adapting the system to the conditions of the dierent surveillance areas.

REFERENCES

[1] M. Shah, O. Javed, and K. Shaque. Automated visual surveillance in realistic scenarios. IEEE Computer Society, 14:30-39, Jan.-March 2007.

[2] A. Hampapur, L. Brown, J. Connell, S. Pankanti, A. Se-

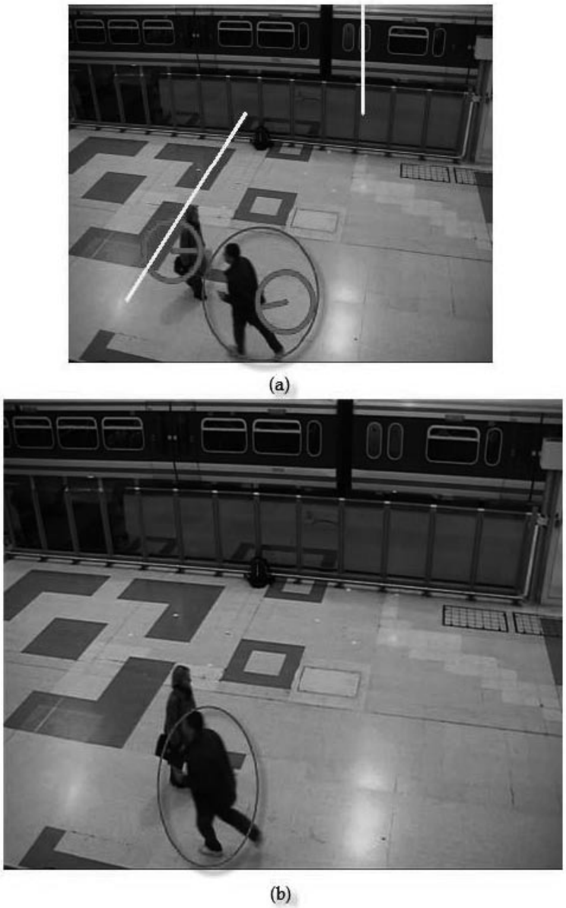

Figure 6. (a) Detected run (360x288). (b) Proof image (720x576).

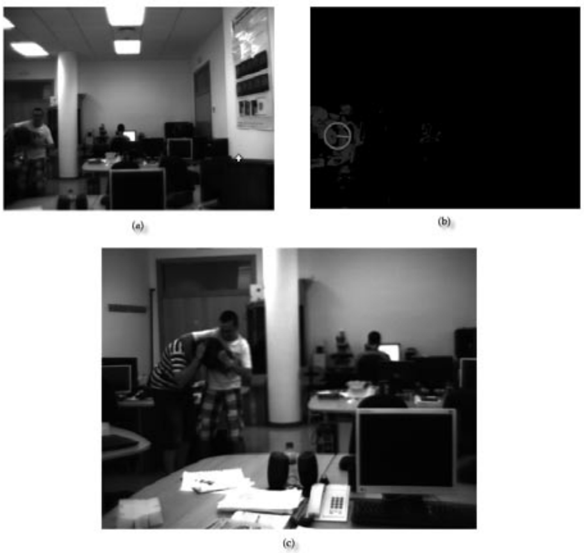

Figure 7. (a) Detected ght (512x384). (b) MHI (512x384).
(c) Stored proof (1024x768).

nior, and Y. Tian. Smart video surveillance: Applications, technologies and implications. In IEEE Pacic-Rim Conference On Multimedia, Singapore, December, 2:1133-1138, 2003.

[3] V. Gouaillier and A. Fleurant. Intelligent video surveillance: Promises and challenges. Technological and Commercial Intelligence Report, March 2009.

[4] Anthony R.Dickand and Michael J.Brooks. Issues in automated visual surveillance. School of Computer Science, University of Adelaide, 2003.

[5] Nathan Johnson. Background subtraction: Various methods for dierent inputs. CiteSeerX Scientic Literature Digital Library and Search Engine [http://citeseerx.ist.psu.edu/oai2] (United States), 2008.

[6] Sen-Ching S. Cheung and Chandrika Kamath. Robust techniques for background subtraction in urban trac video. Video Communications and Signal Based Surveillance, IEEE Conference on, 0:30, 2006.

[12] Gary R. Bradski. Real time face and object tracking as a component of a perceptual user interface. Applications of Computer Vision, IEEE Workshop on, 0:214-219, 1998.

AUTHORS

Enrique Bermejo: Degree in Computer Engineering. Ph.D. Student and associated researcher at Ingeniera de Sistemas y Automatica (ISA) department at E.T.S. Ingenieros Industriales, Universidad de Castilla-La Mancha, Spain.
Enrique.Bermejo@uclm.es

Oscar Deniz: He received his MsC and PhD from Universidad de Las Palmas de Gran Canaria, Spain, in 1999 and 2006, respectively. He has been associate professor at Universidad de Las Palmas de Gran Canaria from 2003 to 2007 and currently at Universidad de Castilla-La Mancha, Spain. His main research interests are human-robot interaction and computer vision. He is a research fellow of the Institute of Intelligent Systems and Numerical Applications in Engineering and mem ber of IEEE, AEPIA and AERFAI. Oscar.Deniz@uclm.es

Gloria Bueno: She received her MsC from Universidad Complutense de Madrid in 1993, and her PhD from Coventry University in 1998. From 1998 to 2000 Gloria worked as a postdoctoral researcher at Universite Louis Pasteur, Strasbourg. In 2000-2001 she worked at CNRS-Institut de Physique Biologique-Hopital Civil and from 2001 to 2003 she was a senior researcher at CEIT (Centro de Estudios e Investigaciones Tecnicas de Gipuzkoa), San Sebastian, Spain. She is currently an Associate Professor at Universidad de Castilla-La Mancha, Spain. Her main research interests include image processing particularly for biomedical engineering applications- computer vision, articial intelligence, modeling and 7 simulation. Gloria.Bueno@uclm.es