# An Innovative Approach for Abandoned or Removed Objects Detection

Tommaso Martiriggiano, Andrea Caroppo, Marco Leo, Paolo Spagnolo, and Tiziana D'Orazio

*Abstract*— **In this paper, a new method is presented for a robust and efficient analysis of video sequences that allows the extraction of foreground objects and the classification of static foreground regions as abandoned or removed objects (ghosts). As a first step, the moving regions in the scene are detected by subtracting to the current frame a background model continuously adapted. Then, a shadow removing algorithm is used to extract the real shape of detected objects and finally, moving objects are classified as abandoned or removed by analysing the boundaries of static foreground regions. The method was successfully tested on real image sequences and it run about 7 fps at size 480x640 on a 2,33 GB Pentium IV machine.**

## I. Introduction

RELIABLE detection of moving objects is an important requirement for video surveillance applications. In these systems, motion detection algorithms can be used to determine the presence of people, cars or other unexpected objects and then start up more complex activity recognition steps.

In the literature, the problem of moving object segmentation is discussed, identifying three different kinds of approaches: optical flow [1, 2], temporal differencing [3], and background subtraction. In particular, methods based on background subtraction, using an opportune threshold procedure on the difference between each image of the sequence and a model image of the background, are recognized by the scientific community as those that provide the best compromise between performance and reliability. Basically, these approaches consist of two steps: the proper updating of a reference background model, and the suitable subtraction between the current image and the background model.

In the past, many approaches based on background subtraction were proposed. Such methods differ mainly in the type of background model and in the procedure used to update the model. In [4, 5] a mixture of Gaussian distributions has been used for modelling the pixel intensities, assuming that more than one process can be observed over time. Pixel values that do not fit the background distributions are considered foreground.

In [6] the authors propose a simple background subtraction method based on logarithmic intensities of pixels. They claim to have results that are superior to traditional difference algorithms and which make the problem of threshold selection less critical. In [7] a prediction-based online method for modeling dynamic scenes is proposed. The approach seems to work well, although it needs a supervised training procedure for the background modeling, and requires hundreds of images without moving objects. Adaptive Kernel density estimation is used in [8] for a motion-based back-ground subtraction algorithm. In this work, the authors use optical flow for the detection of moving objects; in this way, they are able to handle complex background, but the computational costs are relatively high. An interesting approach has been proposed recently in [9]. The authors propose to use spectral, spatial and temporal features, incorporated in a Bayesian framework, to characterize the background appearance at each pixel. Their method seems to work well in the presence of both static and dynamic backgrounds.

Although many researchers focus on the background subtraction, few papers can be found in the literature for foreground analysis [10, 11]. Cucchiara et al. [10] analysed the foreground as moving object, shadow, and ghost by combining the motion information. The computation cost is relatively expensive for real-time video surveillance systems because of the computation of optical flow. In [11] the authors proposed a background subtraction system designed to detect moving objects in a wide variety of conditions, and a second system to detect objects moving in front of moving back-grounds. In this work, a gradient-based method is applied to the static foreground regions to detect the type of the static regions as abandoned or removed objects (ghosts). It does this by analysing the change in the amount of edge energy associated with the boundaries of the static foreground region between the current frame and the background image. By our knowledge, the performance of this method could strongly depend on the technique used to update the background and, moreover, they could fail in presence of non uniform objects.

In this paper, we propose a motion detection system, based on background subtraction algorithm, able to classify static foreground regions as abandoned or removed objects. A template matching procedure is applied between the edge of the foreground region and the edge detected over the segmented image. The rest of the paper is organized as follow: an overview of the proposed system is provided in section II, where motion detection, shadow removing, discrimination between removed versus abandoned objects, will be detailed; finally, section III presents the experimental results obtained on the real image sequences acquired by IEEE 1394 cameras in our laboratory.

## II. SYSTEM OVERVIEW

The proposed system processes the acquired images by a motion detection algorithm performed through background subtraction. In this phase, the background is automatically built and updated by temporal statistical analysis. After motion detection, a shadow removing procedure is performed on each image in order to discard shadow points that, generally, deform the shape of the moving objects.

By analysing the edges, the system is able to detect the type of static regions as abandoned object (a static object left by a person) and removed object (a scene object that is moved).

Following subsections will explain the details of each algorithmic step involved.

### *A. Motion detection*

The implemented motion detection algorithm for moving object extraction is based on background subtraction. It is composed of three distinct phases: firstly, a model of the background needs to be created; then a background subtraction procedure is used to distinguish moving objects from static ones. Finally, an updating algorithm adapts the background to any variation in light conditions.

The background modeling algorithm implemented is very reliable because it does not require any assumption about the presence of moving objects in the scene.

It uses a sliding window (of N frames) whose first frame is assumed as the first coarse background model, even if there are moving objects. Then, each frame of this window is compared with the coarse background: if a pixel value is similar (in all the three color channels) to the correspondent in the model image, mean value and standard deviation are evaluated for that point.

Practically, for each pixel, 6 parameters are considered: $\mu_R, \mu_G, \mu_B, \sigma_R, \sigma_G, \sigma_B$, where $\mu_n$ and $\sigma_n$ represent respectively the mean value and the standard deviation in the n-th color band.

After checking all frames of the examined window, the statistical parameters are maintained only for those pixels with intensity values similar to the model for at least 90% of the whole considered window.

After this, a new sliding window is examined using as referring model the statistical parameters where maintained and the intensity values in the first image for those pixels for which the statistical parameters are rejected in the previous step.

This procedure is iterated until a mean and a standard deviation value have been estimated for all the pixels.

After the model construction, the system is able to automatically detect the presence of moving objects. For this purpose, a simple subtraction algorithm has been implemented. It is based on the evaluation of the difference between current image and the model; this difference is calculated for each color band. A pixel will be considered as a moving point if it differs more than two times from the relative variance at least in one color band. Formally, denoting with $I_{OUT}$ the output binary image:

$$I_{OUT}(x,y) = \begin{cases} 1 \quad \text{if} & |I_R(x,y) - \mu_R(x,y)| > 2 * \sigma_R(x,y) \vee \\ & \vee |I_G(x,y) - \mu_G(x,y)| > 2 * \sigma_G(x,y) \vee \\ & \vee |I_B(x,y) - \mu_B(x,y)| > 2 * \sigma_B(x,y) \\ 0 \quad \text{otherwise} & \end{cases}$$

In order to make the system substantially insensible to variations in light conditions, an updating module has been implemented.

The characteristics of the application context requires some specific constraints: in particular, objects that differ from the background image have always to be detected, that is they will be never included in the background model in order to maintain in-formation about the presence of object removed from the scene until anomalous conditions will be restored.

So, the updating procedure starts from the output of the last algorithm, and only the pixels corresponding to static points ($I_{OUT}(x,y)=0$) will be updated. In detail, for each point, a weighted mean between the historic value and current value is carried out. The parameter a used for the updating can vary in [0,1] and smoothes the relative relevance of the current image instead of the background one

$$\mu_R^{t+1} = \begin{cases} \alpha * \mu_R^t + (1-\alpha) * I_R^t & \text{if} \quad I_{OUT} = 0 \\ \mu_R^t & \text{if} \quad I_{OUT} = 1 \end{cases}$$

### *B. Shadow removing*

After the background subtraction only the blobs whose area is greater than a certain threshold are maintained.

Unfortunately each preserved blob contains not only the relative moving object but also its own shadows. The presence of shadows is a great problem for a motion detection system, because they alter real size and dimension of the objects. This problem is more complex in indoor contexts, where shadows are emphasized by the presence of many reflective objects; in addition shadows can be detected in every direction, on the floor, on the walls but also on the ceiling, so typical shadow removing algorithms, that assume shadows in a plane orthogonal with the human plane, cannot be used.

To prevent all these problems, correct shapes of the objects must be extracted and to do that a shadow removing algorithm is implemented.

The shadow removing approach described here starts from the assumption that a shadow is a uniform decrease of the illumination of a part of an image due to the interposition of an opaque object with respect to a bright point-like illumination source. From this assumption, we can note that shadows move with their own objects but also that they do not have a fixed texture, as real objects do: they are half-transparent regions which retain the representation of the underlying background surface pattern. Therefore, our aim is to examine the parts of the image that have been detected as moving regions from the previous segmentation step but with a texture substantially unchanged with respect to the corresponding background. To do it, we look for moving points whose attenuation values, at each color band, are similar; differently, moving points belonging to true foreground regions will have different attenuation values. In addition, these attenuation value will be lower than 1, because of the minor light that illuminates the shadow regions. Formally, we evaluate, for each moving point (x,y) the attenuation values S at each color band:

$$S_R(x,y) = \frac{I_R(x,y)}{B_R(x,y)} \; S_G(x,y) = \frac{I_G(x,y)}{B_G(x,y)} \; S_B(x,y) = \frac{I_B(x,y)}{B_B(x,y)}$$

where $I_n(x,y)$ and $B_n(x,y)$ are respectively the intensity value in the n-th color band of the pixels (x,y) in the current image and in the background image.

After this, pixels with an uniform attenuation will be removed:

$$I_{OUT}(x,y) = \begin{cases} 0 \;\; if \;\; S_R(x,y) \cong S_G(x,y) \cong S_B(x,y) \wedge \\ \quad \wedge S_R(x,y), S_G(x,y), S_B(x,y) < 1 \\ 1 \quad otherwise \end{cases}$$

The output of this phase provides a motion image with the real shape of the moving objects, without noise or shadows.

### *C. Abandoned and Removed Objects Detection*

In many video surveillance applications is very important to distinguish between abandoned and removed objects.

When a static foreground region is detected, we consider the segmented image (Fig. 1c), after shadow removing step, relative to current frame (Fig. 1b). The next step consists in applying an edge algorithm on the foreground region in the segmented image, obtaining the image in Fig 1e. The same portion is selected in the real image (Fig. 1b) on which the edge algorithm is newly applied (see Fig. 1d).

Now, the two images containing the edges are matched and a similarity measure is calculated. Finally, if this measure is more than a predefined threshold then we decide that an object is abandoned in the scene, otherwise we decide that an object is re-moved from the background.

To perform edge detection, we use Susan algorithm [12], that is very fast and has optimal performances.

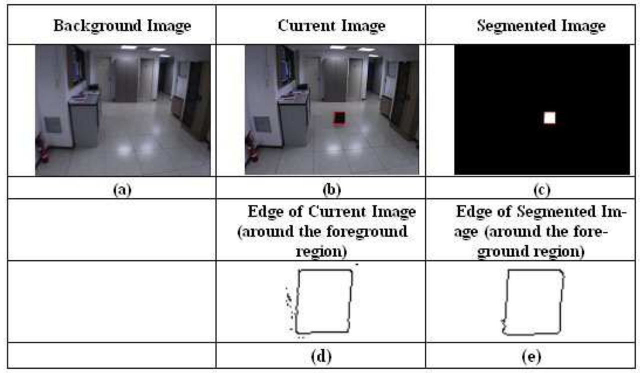



**Fig. 1.** An example of abandoned object in the corridor of a laboratory; (a) background model, (b) current frame with a red rectangle around the detected object, (c) segmented image obtained by the procedure of motion detection and shadow removing. Finally, (d) edges detected in the red rectangle of the current image, (e) edges detected in the red rectangle of the segmented image.

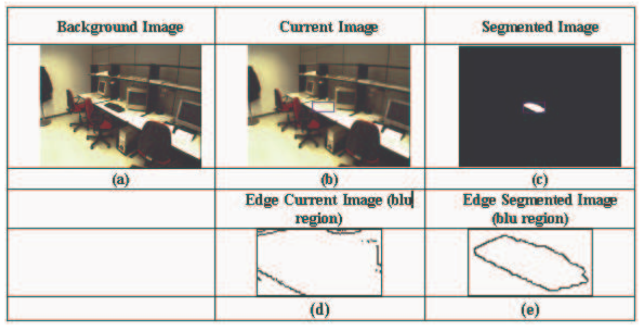



**Fig. 2.** An example of a removed object in a room of the laboratory: (a) model of background, (b) current frame with a blue rectangle around to the region of removed object, (c) segmented image obtained by the procedure of motion detection and shadow removing, (d) edges detected in the blue rectangle of the current image, (e) edges detected in the blue rectangle of the segmented image.

## III. EXPERIMENTAL RESULTS

In this section, some experiments, performed in a laboratory, demonstrate the effectiveness of the method proposed. The algorithm runs about 7 fps for color images at size 480x640 on a 2,33 GB P IV PC. Following subsections will show the results of each algorithmic step.

Firstly, we only applied the motion detection algorithm on the original image, obtaining the results shown in the second column of the Fig.3. We notice that the real shape of moving persons is largely modified due the presence of shadows, moreover, in some cases, there is only one foreground segmented region produced by two moving persons. This kinds of problems have been resolved by our shadow suppression algorithm in a very good way as it can be seen in the third column of the Fig. 3.

The decision between removed/abandoned object has been taken by the new technique that we have introduced in subsection II.B. This algorithm is based on a template matching procedure that compute a similarity measure between the edge detected on the foreground region and the edge detected over the segmented image. Therefore, the decision between abandoned or removed objects is taken comparing the obtained similarity measure with a established threshold value.

**Fig. 3.** The figure shows the results obtained by motion detection and shadow removing algorithms on images acquired in a laboratory. The first column shows the original image, the second column shows the segmented image obtained by only motion detection algorithm, the third column shows the image obtained by motion detection and shadow removing algorithms.

They have been carried out two experiments, both in a room of the laboratory. The obtained experimental results are visualized in the next two tables.

**Table 1.** First experiment in the laboratory: a bag is abandoned on the desktop

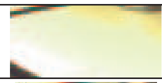
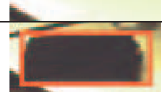
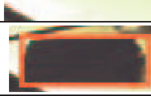
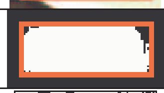
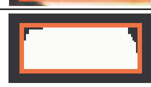
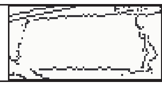
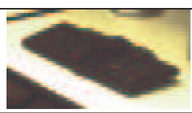
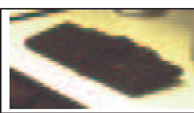
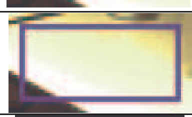
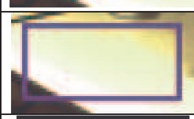
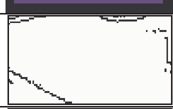
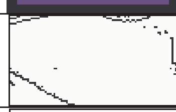
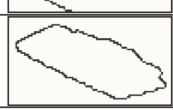
| | | |
|---|---|---|
| **Background Image** |  |  |
| **Current Image** |  |  |
| **Segmented Image** |  |  |
| **Edge Current Image** |  |  |
| **Edge Segmented Image** |  |  |
| **Matching %** | 74 % | 69 % |

**Table 2.** Second experiment in the laboratory: a keyboard is removed from the desktop

| | | |
|---|---|---|
| **Background Image** |  |  |
| **Current Image** |  |  |
| **Segmented Image** |  |  |
| **Edge Current Image** |  |  |
| **Edge Segmented Image** |  |  |
| **Matching %** | 6 % | 5 % |

In the last row of the tables are reported the matching percentages; generally, in our tests, we decide for an abandoned object if the matching percentage is more than 65% and we have labelled the object with a red rectangle; on the other hand, we established that an object was removed from the background if the matching percentage is less than 30% and we have labelled the region with a blue rectangle. If the percentage is comprised between 30% and 65%, the algorithm is not able to take a decision. As shown in the tables, our procedure was able to correctly classify the situations of removed/abandoned objects in all experiments. Finally, we note that when an object is abandoned the matching percentage is very high, while when the object is removed we obtain very low matching values; this demonstrates the robustness of the algorithm, since the choice of the threshold is not critical.

## IV. CONCLUSIONS

In this work, we proposed a new method to efficiently analyse foreground. As a first step, an adaptive background model on the RGB images acquired by common digital cameras has been implemented. After the detection of moving regions, a shadow re-moving algorithm has been implemented in order to clean the real shape of the detected objects. Finally, we discriminate between abandoned or removed objects by analysing the boundaries of static foreground regions. The reliability of the proposed framework is shown by large experimental tests performed in our laboratory.

## REFERENCES


[1] Fejes, S., Davis, L.S.: Detection of independent motion using directional motion estimation. Technical Report. CAR-TR-866, CS-TR 3815. University of Maryland (1997)

[2] Fejes, S., Davis, L.S.: What can projections of flow fields tell us about the visual motion. ICCV'98, 4-7 Jan. Bombay, India

[3] Paragios, N., Deriche, R.: Geodesic Active Contours and Level Sets for the Detection and Tracking of Moving Objects., IEEE Trans. On PAMI, Vol.22, 3 pp. 266-280 (2000)

[4] Grimson, E., Stauffer, C., Roman, R., Lee, L.: Using adaptive tracking to classify and monitoring activities in a site. CVPR 98, pp.22-29, Santa Barbara, CA, USA, 23-25 June (1998)

[5] Stauffer, C., Grimson, W., E., L.: Learning patterns of activity using real-time tracking. IEEE Transaction on Pattern Analysis and Machine Intelligence Vol.22 N. 8, pp. 747-757, (2000)

[6] Quen-Zong Wu, Bor-Shenn Jeng: Background subtraction based on logarithimc intensities. Patter Recognition Letters 23, pp. 1529-1536, (2002)

[7] Monnet, A., Mittal, A., Paragios, N., Ramesh, V.: Background Modeling and Subtraction of Dynamic Scenes. Proc. of Int. Conference on Computer Vision, pp 1305-1312, (2003)

[8] Mittal, A., Paragios, N.: Motion-based Background Subtraction using adaptive kernel Density Estimation. Proc. of CVPR, pp 302- 309, (2004)

[9] Li, L., Huang, W., Gu, I.,Y.,H., Tian, Q.: Statistical Modeling of Complex Backgrounds for Foreground Object Detection. IEEE Trans. on Image Processing, Vol. 13, No. 11, (2004)

[10] Cucchiara, R., Grana, C., Piccardi, M., Prati, A.: Detecting Moving Objects, Ghosts, and Shadows in Video Streams. IEEE Trans. on PAMI, 25: (10), October, (2003)

[11] Connell, J.: Detection and Tracking in the IBM People Vision System. IEEE ICME, (2004)

[12] Smith., S.M.: A new class of corner finder. Proc. 3rd British Machine Vision Conference, pages 139-148, (1992)