To appear in IET Electronic Letters

Pixel-based colour contrast for abandoned and stolen object discrimination in video surveillance

J.C. SanMiguel, L. Caro and J.M. Martinez

A novel approach is proposed for discriminating between abandoned or stolen previously detected stationary foreground regions in video surveillance. It is based on measuring the colour contrast of the contour of the stationary object under analysis at pixel level. Two contrasts are computed by analyzing such contour in the current and background frames. Then, both are combined for performing the discrimination. The experimental results over a heterogeneous dataset containing real scenarios demonstrate that our approach outperforms the related literature and highly reduces the computational cost of the discrimination task allowing real-time operation.

Introduction: In the video surveillance domain, the automatic detection of abandoned and stolen objects in real-time has recently become a topic of great interest especially in crowded environments [1]. In general, this detection is achieved by developing a system with the following analysis stages: foreground segmentation, stationary region detection, blob classification and abandoned/stolen discrimination [1]. The last stage of this pipeline determines the system ability to discriminate stationary foreground objects between abandoned and stolen. For its implementation, the common approach is to study the similarities between features extracted from the current and background frames of the video sequence. The related literature can be classified depending on the feature employed: edge, colour, contour or hybrid. Edge-based approaches accumulate the gradient energy at pixel level along the object contour in both frames [2]. Colour-based approaches assume that the surrounding region to the object is very similar to the background that has been covered (abandoned case) or uncovered (stolen case) and, for the similarity, they use colour histograms at region-level [3]. Hybrid approaches combine colour and edge to improve the discrimination process [4]. The main limitation of the edge- and colour-based approaches is that they need homogeneous properties in the regions of the background close to the stationary object (in terms of colour, motion and edges) and rely on precise foreground segmentation masks. Therefore, their accuracy is reduced in complex situations. Recently, contour-based approaches have been proposed to increase the robustness in complex situations by applying adjustments of the object contour (e.g., active contours) [5]. However, its iterative nature restricts the real-time operation of the final video surveillance system.

This Letter presents a new approach for discriminating stationary objects into abandoned and stolen by using the colour contrast along the object contour at pixel level. It assumes that object contour coincides with the colour boundaries of the frame. Opposed to current approaches, it does not require specific background properties being suitable for complex backgrounds and non-accurate foreground segmentation masks allowing real-time operation.

Discrimination scheme: The block diagram of the proposed discrimination scheme is depicted in Fig. 1. It starts from the initial contour of the stationary object at time t, defined as the set of points $C_t = \{p_1, ..., p_i, ...p_N\}$. where p_i represents the x,y coordinates of the ith pixel of this contour and N is the total number of points. In our approach, the contour extraction consists on point-scanning the result of applying the Canny edge detector to the current foreground mask. Then, the average contrast between points inside and outside the detected region is computed, on the current frame and the background images. These contrast measures are computed as follows:

$$A_{PCC}^F = z(F_t, C_t), \quad (1)$$

$$A_{PCC}^{B} = z(B_t, C_t) \, , \tag{2}$$

where $z(\cdot)$ denotes the technique for contrast analysis; and F_t are B_t the current and background frames; and A_{PCC}^{F} and A_{PCC}^{B} are the colour contrast results in those frames, respectively. For increasing robustness, colour information from all channels is employed to compute the averages. Finally, the two average measures are subtracted to generate a final score, S_{PCC} , as follows:

$$S_{PCC} = A_{PCC}^{B} - A_{PCC}^{F}. \tag{3}$$

Finally, a threshold th is used to decide whether the object is abandoned or stolen.

Pixel colour contrast: For computing such contrast, we use the spatial boundary contrast metric proposed in [6]. For each pixel p_i of the contour, segments of length $2L+1$, normal to the contour's curve, are defined. The values of the pixels on both ends of the segment, namely points P_I and P_O, are then compared. This comparison is performed by defining a small window of size $M \times M$ centred in those pixels. This scheme is illustrated in Fig. 2. The distance measure between the two endpoints, Boundary Spatial Color Contrast (BSCC), is defined for each boundary pixel as follows:

$$BSCC(F_t, C_t, i) = \frac{\left\| W_O^i(t) - W_I^i(t) \right\|}{\sqrt{3 \times 255^2}}, \tag{4}$$

where W_O and W_I are the average colour values computed in the $M \times M$ neighbourhood of points P_O and P_I (in the RGB colour space) for the ith contour pixel of the C_t in the F frame (that could be either the current or background frames). This measure is only defined for those boundaries pixels for which P_O , P_I

and the pixels inside their neighbourhoods fall inside the image boundaries (considering as non-valid those pixels that fall outside image boundaries). Then, the average BSCC value along the analyzed contour pixels is then expressed as follows:

$$z(F, C_t) = \frac{1}{K_t} \sum_{i=1}^{K_t} BSSC(F, C_t, i), \qquad (5)$$

where K_t is the total number of analyzed pixels with valid values, and BSCC is the spatial color contrast measure for the ith pixel. This function computed on both the current frame and the background (as shown in equations (1) and (2)), and combined to obtain the detector's score (as shown in equation (3)). A_{PCC}^{B} is expected to have a value close to 0.0 for abandoned objects, and a higher value for stolen objects, due to the contrast between the object and its surroundings; with A_{PCC}^{F} getting opposite values in the same situations.

Experimental results: We have evaluated the proposed approach using the ASODds dataset [5]. In particular, we have used the real data that consists on foreground masks representing the stationary objects of the scene which contains three categories with increasing complexity. In addition, we have compared our proposal against three representative approaches based on edge (ED [2]), colour (CH[3]) and contour (CO[5]). For our approach, the threshold th was learned by means of ROC analysis. The metrics used for comparison are the correct classification accuracy, defined as the ration between the correct and the total number of discriminations, and the computational cost of the task. Accuracy and time results are reported in, respectively, Table 1 and 2. The accuracy results indicate that the proposed approach achieves higher performance than the ED and CH approaches and slightly better than the CO approach. The use of real data (that is inaccurate) with varying complexity demonstrates the robustness of the approach. Computational cost results show that our approach highly reduces the time execution. A reduction factor higher than 92% was achieved.

Conclusion: This Letter proposes a novel approach for discriminating stationary foreground objects into abandoned or stolen based on boundary colour contrast at pixel level. It assumes the matching between the object contour and the colour boundaries of the video frame. As it can be observed in the experimental results, our approach improves the accuracy of related literature in video sequences with varying complexity and highly reduces the computational complexity of the discrimination task. Hence, our approach is suitable for developing real-time applications in complex environments. In addition, a potential application of the proposed approach is the use of the pixel colour contrast for detecting and filtering common artifacts generated by the background subtraction technique such as ghosts or incorrectly segmented regions

Acknowledgments: This work has been partially supported by the Spanish Government (TEC2011-25995), by the Consejería de Educación of the Comunidad de Madrid and by The European Social Fund.

References

1 TIAN, Y., FERIS, R., LUI, H., HUMPAPUR, A., and SUN, M.: 'Robust detection of abandoned and removed objects in complex surveillance videos', IEEE Trans. on Systems, Man, and Cybernetics (Part C), 2010, Vol. 41(5), pp. 565-576

2 SPAGNOLO, P., CAROPPO, A., LEO, M., MARTIRIGGIANO, T., and D'ORAZIO, T.: 'An Abandoned/Removed Objects Detection Algorithm and Its Evaluation on PETS Datasets', Proc. of Int. Conf. on Advanced Video and Signal based Surveillance, Sidney, Australia, 2006, Vol. 1, pp. 17-21

3 FERRANDO, S., GERA, G. and REGAZZONI, C: Classification of unattended and stolen objects in videosurveillance system. Proc. of Int. Conf. on Advanced Video and Signal Based Surveillance, Sidney, Australia, 2006, Vol. 1, pp. 21-27

4 WEN, J., GONG, H., ZHANG, X., and HU, W.: 'Generative Model for Abandoned Object Detection', Proc. of Int. Conf. on Image Processing, Cairo, Egypt, 2009, Vol. 1, pp. 853-856

5 CARO, L., SANMIGUEL, J.C., and MARTINEZ, J.M. : 'Discrimination of abandoned and stolen object based on active contours', Proc. of Int. Conf. on Advanced Video and Signal based Surveillance, Klagenfurt, Austria, 2011, Vol. 1, pp. 101–106

6 ERDEM, C.E., SANKUR, B. and TEKALP, M.: 'Performance measures for video object segmentation and tracking'. IEEE Trans. on Image Processing, 2004, Vol. 13(7), pp. 937-951

.**Authors' affiliations:**

J.C. SanMiguel (corresponding author), L. Caro and J.M. Martínez
(Video Processing and Understanding Lab, Escuela Politécnica Superior,
Universidad Autónoma de Madrid, 28049, Madrid, Spain)
E-mail: juancarlos.sanmiguel@uam.es

Figure captions:

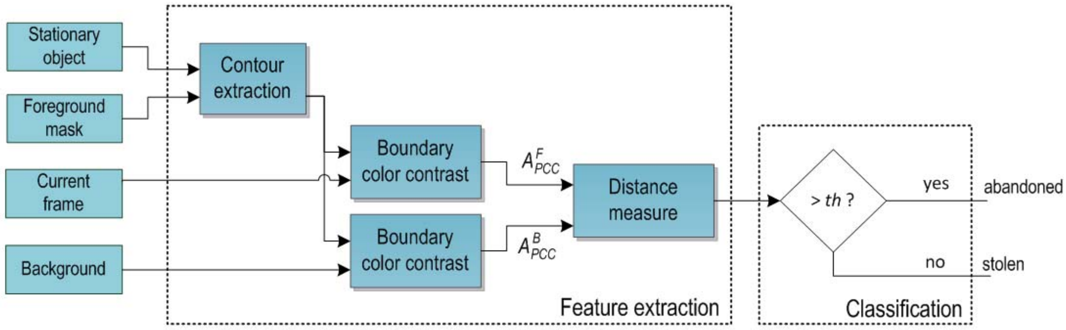
Fig. 1 Proposed scheme for abandoned and stolen object discrimination.

Fig. 2 Pixel color contrast detector: (a) static foreground object, (b) analyzed points along the boundary and (c) analyzed contour point.

Table 1 Correct classification accuracy for the discrimination task (mean ± standard deviation).

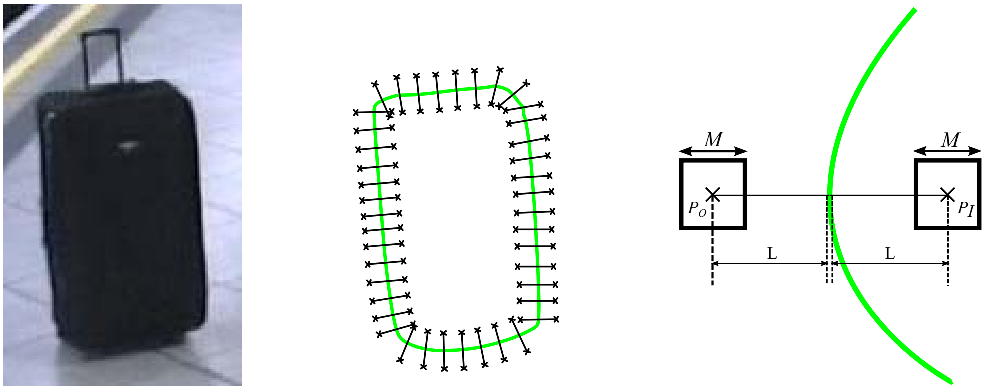
Table 2 Computational cost for the discrimination task (ms).

Figure 1



Stationary object
Foreground mask
Current frame
Background
Contour extraction
Boundary color contrast
Boundary color contrast
A^F_{PCC}
A^B_{PCC}
Distance measure
Feature extraction
> th ?
yes
abandoned
no
stolen
Classification

Figure 2



(a) (b) (c)

Table 1

Approach	Category 1	Category 2	Category 3	Mean
CH	.755±.035	.712±.055	.848±.025	.777±.024
ED	.918±.024	.791±.039	.743±.041	.821±.022
CO	.960±.010	.952±.027	.929±.016	.947±.011
Proposed	.967±.014	.943±.013	.951±.013	.954±.009

Table 2

Approach	Minimum	Maximum	Mean
CH	5.6	44.5	23.2
ED	0.1	133.8	28.3
CO	20.7	1187.1	246.3
Proposed	0.2	8.6	1.9