

Moving Object Detection for Real-Time High-Quality Lightweight Applications on Smart Cameras

Carlos Cuevas, Narciso García
Universidad Politécnica de Madrid, Spain

Abstract— Here, a novel and efficient strategy for moving object detection by non-parametric modeling on smart cameras is presented. Whereas the background is modeled using only color information, the foreground model combines color and spatial information. The application of a particle filter allows the update of the spatial information and provides a priori information about the areas to analyze in the following images, enabling an important reduction in the computational requirements and improving the segmentation results.

I. INTRODUCTION

Recently, the increased computational speed of processors has enabled the use of new computer vision applications in several fields such as video-surveillance, monitoring, or human-machine interaction [1]. These applications include, as a key step for high level analysis tasks, a moving object detection strategy. As a result, a large number of moving object segmentation approaches has been proposed in the literature. Some of these algorithms aim to maximize the speed and to reduce the memory requirements, but they are not efficient in complex situations such as noisy sequences, illumination changes or dynamic backgrounds [2].

To tackle these limitations, several multimodal strategies have been proposed. Among them, non-parametric methods have been developed to solve the problems in environments where the pixel variations cannot be described parametrically [3]. However, their main drawback is that they result in very high memory and computational requirements.

Here, we propose a novel and very fast non-parametric segmentation strategy. Using a particle filter based tracking strategy over previously detected foreground regions, the segmentation is improved and the computational cost is reduced. Moreover, the particle filter provides probabilistic information about the areas where the moving objects are expected to appear in the following images. Applying the non-parametric modeling exclusively over these areas, combined with the result of a Windowed Random Sampling (WRS), an important reduction in computational requirements is achieved. Additionally, unlike previously segmentation strategies, the spatial information is considered only for the foreground modeling, resulting in an additional cutback of the computational and memory necessities.

The proposed strategy can be applied to smart cameras operating in real-time either in outdoor or in indoor scenarios.

This work has been partially supported by Alcatel-Lucent and the Spanish Administration agency CDTI under project CENIT 2007-1012 i3media and by the Ministerio de Ciencia e Innovación of the Spanish Government under project TEC2007-67764 (SmartVision).

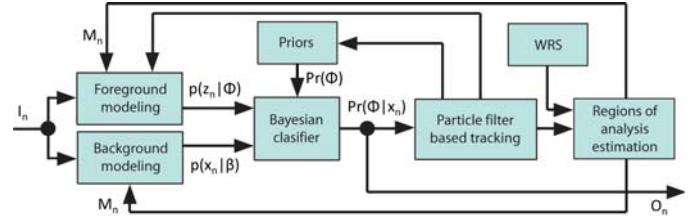


Fig. 1. Block diagram of the proposed system.

II. SYSTEM DESCRIPTION

The proposed strategy, depicted in Figure 1, contains two main steps: a Bayesian classifier, where the moving objects are separated from the background; and a tracking strategy, which improves the efficiency of the Bayesian modeling and reduces the area to be analyzed in the following image.

A. Non-parametric modeling

Let us consider, for a pixel x_n in the current image, the set of color samples $x_i = (R_i, G_i, B_i)$ from the N_β previous images. Applying Gaussian kernels with diagonal bandwidth matrices, the background likelihood for x_n is

$$p(x_n | \beta) = \frac{1}{N_\beta (2\pi)^{3/2}} \sum_{i=1}^{N_\beta} \prod_{j=1}^3 \frac{1}{|h_j|} \exp\left(-\frac{(x_n(j) - x_i(j))^2}{2h_j^2}\right) \quad (1)$$

where β is the background class and h_j is the bandwidth for each color component.

Taking into consideration that foreground objects are in motion, the foreground likelihood should be constructed from samples defined as $z_i = (R_i, G_i, B_i, r_i, c_i)$, where (r_i, c_i) are the spatial coordinates. Then, the foreground likelihood for a current pixel z_n , can be estimated as [3]

$$p(z_n | \Phi) = \alpha\gamma + \frac{(1-\alpha)}{M (2\pi)^{5/2}} \sum_{i=1}^M \prod_{j=1}^5 \frac{1}{|h'_j|} \exp\left(-\frac{(z_n(j) - z_i(j))^2}{2(h'_j)^2}\right) \quad (2)$$

where Φ is the foreground class and h'_j is the bandwidth for each foreground component.

Once we have modeled both background and foreground, using Bayes' theorem we can write

$$\Pr(\Phi | x_n) = \frac{\Pr(\Phi) p(z_n | \Phi)}{\Pr(\Phi) p(z_n | \Phi) + \Pr(\beta) p(x_n | \beta) p(r_n, c_n | \Phi)} \quad (3)$$

where $\Pr(\Phi)$ and $\Pr(\beta)$ are the prior probabilities for foreground and background respectively, and $p(r_n, c_n | \Phi)$ is the marginalization of the estimated foreground likelihood, $p(z_n | \Phi)$, over the color components [4].

B. Tracking strategy

Among others, the main difference between previous works and the proposed here is the update of the foreground spatial information through a multi-region tracking strategy.

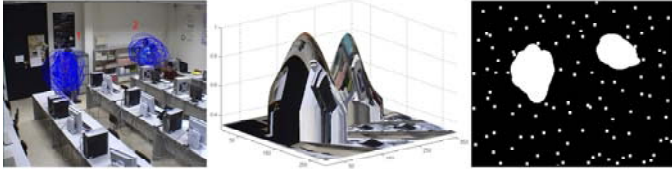


Fig. 2. Some results from the particle filter. Left, original image with two moving objects and their corresponding predicted particles. Middle, foreground prior values with the projection of the original image. Right, mask resulting from the combination of the WRS and the predicted particles.

This update allows preservation of the required quality by selecting a smaller buffer of images, N_ϕ , and a reduced foreground spatial bandwidth, h_ϕ , resulting in a significant reduction of the computational and memory requirements.

For this purpose we have used a multi-region particle filter based on the filter proposed in [5], which is able to deal with appearing and disappearing foreground regions with no assumptions on relationships between them.

Furthermore, we also obtain the prior information used in (3) from the predicted particles. Moreover, the position of these particles is combined with the result of a Windowed Random Sampling (WRS) to generate a mask, M_n , with the regions to be analyzed in the following image. While the positions of the particles determine the areas where the existing moving objects are expected to appear in the future, the WRS avoid the possible misdetections of new moving objects. An example is presented in Figure 2.

III. EXPERIMENTAL RESULTS

To evaluate the proposed strategy, both the quality of the results and the computational requirements of the system have been analyzed. Therefore, the system has been tested in several indoor and outdoor environments, containing critical situations such as dynamic backgrounds and illumination changes.

TABLE I
COMPUTATIONAL COST AT PIXEL LEVEL

| | Background | | Foreground | |
|---------------------------|-----------------------------|--------|-----------------------------|--------|
| | Cost | Ratio | Cost | Ratio |
| Strategies in [3] and [6] | $\propto h_\beta^2 N_\beta$ | 10^3 | $\propto h_\phi'^2 N_\phi'$ | 10^2 |
| Ours strategy | $\propto N_\beta$ | 1 | $\propto h_\phi^2 N_\phi$ | 1 |

Table I presents, applying typical values [3], the computational comparison for one pixel between our method and two outstanding segmentation strategies based on non-parametric modeling. The modeling of the background without spatial bandwidth, h_β , allows us to reduce three orders of magnitude the number of operations. On the other hand, updating the foreground coordinates, the number of previous images, N_ϕ' , and the foreground spatial bandwidth, h_ϕ' , can be reduced to N_ϕ and h_ϕ , maintaining the required quality and reducing two orders of magnitude the number of operations.

Figure 3 presents some of the obtained results over two representative video sequences. The first one (first row in the figure) presents a complex indoor scenario with two moving

objects. The second one (second row) depicts an outdoor environment with multiple dynamic background objects.

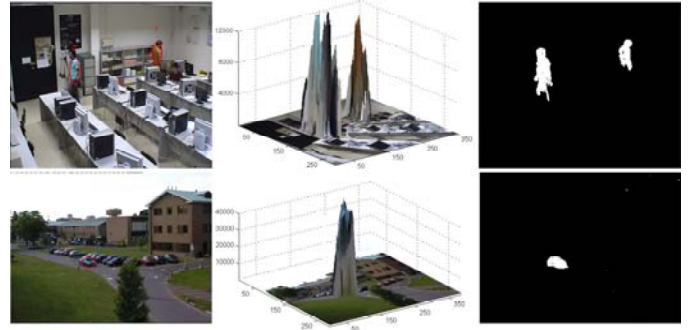


Fig. 3. Results over two different sequences. First column, original images with the state vector estimations. Second column, foreground probability density function. Third column, final segmentations.

Additionally to the high computational reduction at pixel level presented in Table I, the use of M_n reduces drastically the number of pixels to take into account, resulting in a very significant reduction in the memory and computational requirements: a 93.1 % for the indoor sequence of Figure 1, and a 90.4 % for the outdoor sequence.

IV. CONCLUSIONS

A novel and fast high-quality moving object detection strategy for lightweight applications on smart cameras has been presented. While the background is constructed using exclusively color information, the foreground is modeled combining color and spatial information, thus reducing several orders of magnitude the computational requirements.

The proposed tracking strategy, based on a particle filter, updates the coordinates of previously detected foreground regions and therefore improves the segmentation results. It also provides a priori information about the areas where moving objects are expected to appear, achieving an additional reduction of the computational and memory requirements, and improving the performance of the Bayesian classifier.

REFERENCES

- [1] L. Renjie, Y. Songyu, and Y. Xiaokang, "Efficient Spatio-temporal Segmentation for Extracting Moving Objects in Video Sequences," *IEEE Trans. Consumer Electronics*, vol. 53, no. 3, pp. 1161-1167, 2007.
- [2] A. Tavakkoli, M. Nicolescu, G. Bebis, and M. Nicolescu, "Non-parametric statistical background modeling for efficient foreground region detection," *Machine Vision and Applications*, vol. 20, no. 6, pp. 395-409, 2009.
- [3] Y. Sheikh and M. Shah, "Bayesian modeling of dynamic scenes for object detection," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 27, no. 11, pp. 1778-1792, 2005.
- [4] C. Cuevas, R. Moledano, F. Jaureguizar, and N. García, "High-quality real-time moving object detection by non-parametric segmentation," *Electronics Letters*, vol. 46, no. 13, pp. 910-911, 2010.
- [5] M. Nieto, C. Cuevas, and L. Salgado, "Measurement-based reclustering for multiple object tracking with particle filters," in *IEEE Int. Conf. Image Processing*, pp. 4097-4100, 2009.
- [6] X. Zhang and J. Yang, "Foreground segmentation based on selective foreground model," *Electronics Letters*, vol. 44, no. 14, pp. 851, 2008.