AUTOMATIC BANDWIDTH ESTIMATION STRATEGY FOR HIGH-QUALITY NON-PARAMETRIC MODELING BASED MOVING OBJECT DETECTION

Carlos Cuevas and Narciso García

Grupo de Tratamiento de Imágenes - E.T.S. Ing. Telecomunicación
Universidad Politécnica de Madrid - Madrid - Spain
{ccr, narciso}@gti.ssr.upm.es

ABSTRACT

Here, a novel and efficient moving object detection strategy by non-parametric modeling is presented. Whereas the foreground is modeled by combining color and spatial information, the background model is constructed exclusively with color information, thus resulting in a great reduction of the computational and memory requirements. The estimation of the background and foreground covariance matrices, allows us to obtain compact moving regions while the number of false detections is reduced. Additionally, the application of a tracking strategy provides a priori knowledge about the spatial position of the moving objects, which improves the performance of the Bayesian classifier.

Index Terms— Object detection, non-parametric modeling, bandwidth estimation, particle filter, tracking, Mean-Shift.

1. INTRODUCTION

The detection of unusual motion is a key step for high level object analysis tasks such as tracking, classification or event analysis. To achieve very high sensitivity in the detection of moving objects, looking for the lowest possible amount of false alarms, background subtraction techniques are commonly applied. The purpose of these techniques is to efficiently estimate a background model from a sequence of images, and their quality is evaluated according to their speed, memory requirements and accuracy in the results [1].

Throughout the last years, many multimodal strategies have been developed which, modeling multiple states for each pixel, are able to obtain high quality results [2]. Whereas these techniques solve many complex situations with dynamic backgrounds, they do not provide satisfactory results in environments where the pixel variations can not be described parametrically.

To correctly model the pixel variations in complex multimodal scenarios, non-parametric modeling strategies, which make use of Kernel Density Estimators (KDE), have been proposed by several authors in the recent literature [1] [3] [4]. These strategies estimate the probability density function (pdf) of each image pixel from their recent history.

Although these techniques improve the quality of the results in situations with dynamic backgrounds or illuminations changes, sometimes they are not sufficient to distinguish between foreground and background [5]. To address this limitation, some authors combine a foreground modeling with the background pdf, improving the quality of the results in these situations [4].

Nevertheless, these strategies have an important drawback: for each pixel in every frame, the average of a very large amount of multidimensional kernels should be computed, resulting in very high memory and computational requirements. Moreover, an additional drawback of these strategies is that, when the foreground is modeled jointly to the background, the covariance matrices of the kernels are set as fixed matrices [1]. Therefore, looking for a compromise between preserve the multimodality and the presence of noise in the detection, different covariance values should be manually selected depending on the characteristics of each sequence.

Here, we present a novel and efficient non-parametric segmentation strategy. Whereas the foreground is modeled using color and spatial information, the background likelihoods are obtained using exclusively color information, which results in a great reduction of the computational and memory needs. The covariance matrices that determine the "width" of the kernels in the background and foreground modeling processes are dynamically estimated, then preserving the multimodality while reducing the amount of false negatives in the detection. Moreover, a particle filter based tracking strategy, applied over the previously detected foreground regions, improves the quality of the results and provides probabilistic information about those areas where the moving objects are expected to appear in the following images. To obtain the final foreground probability, the foreground and background likelihoods and this a priori information are combined within a novel Bayesian approach.

2. KERNEL BASED DENSITY ESTIMATION

Let $\{\mathbf{x}_i\}_{i=1}^N$ be a set of d -dimensional samples corresponding to the recent history of a pixel. Using the kernel estimator K , the pdf that this pixel will have intensity value $\mathbf{x}$ can be estimated as:

$$\hat{f}(\mathbf{x}) = \frac{1}{N} \sum_{i=1}^N \|H\|^{-\frac{1}{2}} K\left(H^{-\frac{1}{2}}(\mathbf{x} - \mathbf{x}_i)\right) \quad (1)$$

where H is a $d \times d$ positive definite symmetric covariance matrix [1] that specifies the "width" of the kernel K .

The selected covariance matrix is very important for kernel density estimation [6] and numerous approaches have been proposed in the literature with different alternatives. Small bandwidth values are more appropriate in regions of high density, enabling a more accurate estimation of the density in those regions. Nevertheless, if these values are too small, the likelihood estimation can show spurious features. On the other hand, larger bandwidth values are more appropriate in low density areas where few sample points are available. However, selecting too large bandwidth values, the multimodality can be lost.

This work has been partially supported by the Ministerio de Ciencia e Innovación of the Spanish Government under project TEC2010-20412 (Enhanced 3DTV).

Fig. 3. I: Original images. GT: Ground truth. (a) Results modeling exclusively the background. (b) Results combining foreground and background likelihoods. (c) Results by applying the proposed strategy.

Figure 3 shows some of the obtained results. The first column corresponds to an indoor sequence where the most critical aspect is the similarity between the moving object and the background. In the other columns, different outdoor scenarios, with different number of moving objects and multiple non-static background elements, are presented. Applying the background modeling based strategy, (a), the foreground regions are not correctly detected. However, combining a foreground modeling with the background model, (b), the detected moving objects are more compact and accurate. Nevertheless, as this strategy does not estimate the covariance matrices, the amount of false detections has been increased. The obtained results show that, with the proposed algorithm, (c), the moving objects are correctly detected (as we are combining background and foreground modelings) and the amount of false detections has been reduced (as we are estimating the covariance matrices).

Finally, Fig. 4 depicts some Recall and Precision percentages [5], which correspond to the examples in Fig. 3. Here it is possible to appreciate that, using the proposed strategy, we are able to obtain the best compromise for these quality parameters: correctly detected moving objects, while a low amount of false detections.

7. CONCLUSIONS

A novel and computationally efficient background-foreground non-parametric classification strategy has been presented. Whereas the foreground likelihood is constructed combining color and spatial information, the background model is obtained using exclusively color information, thus reducing several orders of magnitude the computational requirements.

To improve the quality of the results, obtaining compact and accurate detections, while reducing the amount of false negatives, we have dynamically estimated the covariance matrices that determine the appropriate 'width' of the kernels in the computation of the background and foreground likelihoods.

Additionally, through a proposed particle filter based tracking strategy, the spatial position of the previously detected foreground regions are updated, which provides a priori knowledge about the

Fig. 4. Recall and Precision percentages.

areas where the moving objects are expected to appear, improving the performance of the Bayesian classifier.

The obtained results show that the proposed strategy provides high quality results in a large amount of complex situations with dynamic backgrounds. In addition, our approach improves the performance of the detections with respect to other non-parametric modeling based strategies in terms of both Recall and Precision percentages.

8. REFERENCES

- [1] 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.
- [2] C. Cuevas and N. García, "Tracking-based non-parametric background-foreground classification in a chromaticity-gradient space," in *IEEE Int. Conf. Image Processing*, 2010, pp. 845–848.
- [3] N. Martel-Brisson and A. Zaccarin, "Unsupervised approach for building non-parametric background and foreground models of scenes with significant foreground activity," in *ACM workshop on Vision networks for behavior analysis*, 2008, pp. 93–100.
- [4] X. Zhang and J. Yang, "Foreground segmentation based on selective foreground model," *Electronics Letters*, vol. 44, pp. 851, 2008.
- [5] 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.
- [6] Q. Wan and Y. Wang, "Background subtraction based on adaptive non-parametric model," in *Intelligent Control and Automation, 2008. WCICA 2008.*, 2008, pp. 5960–5965.
- [7] A. Mittal and N. Paragios, "Motion-based background subtraction using adaptive kernel density estimation," in *IEEE Conf. Computer Vision and Pattern Recognition*, 2004, vol. 2, pp. 302–309.
- [8] A. Elgammal, R. Duraiswami, D. Harwood, and L.S. Davis, "Background and foreground modeling using nonparametric kernel density estimation for visual surveillance," *Proceedings of the IEEE*, vol. 90, no. 7, pp. 1151–1163, 2002.
- [9] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 603–619, 2002.
- [10] M. Nieto, C. Cuevas, and L. Salgado, "Measurement-based reclustering for multiple object tracking with particle filters," in *IEEE Int. Conf. Image Processing*, 2009, pp. 4097–4100.