CONTRIBUTIONS TO THE SEGMENTATION OF MOVING OBJECTS IN VIDEO SEQUENCES

Author:

Maria Isabel de Castro Lopes Martins Pinto Ferreira

Thesis Advisors:

José Luís Alba Castro, Associate Professor, EET, U.Vigo, Spain

Luís Corte-Real, Associate Professor, FEUP, U.Porto, Portugal

Motivation of the work

Unsupervised segmentation of moving objects in video sequences, by **background subtraction (BS):**

- fundamental step in many vision systems and critical factor for the success of the overall system,
- **hard and challenging task**.

Complex and challenging scenarios include poor lighting conditions, sudden illumination changes, nighttime videos, shadows, parasitic background motion, ...

State-of-the-art methods [1]:

- no method has been able to fully deal with all challenges,
- most widely used methods build a statistical model of background pixels [2],
- methods appear to be complementary in nature [3],
- better performance usually at the cost of significant increase in complexity and computational load.

Comparison of methods:

- most datasets do not contain a balanced set of videos presenting real application challenges,
- hard or impossible to compare results computed on different datasets,
- metrics used to evaluate the average performance do not reveal performance frame by frame.

Research Plan (work started while PhD student in the PhD Program in STC, U.Vigo)

1st Year

- ✓ State-of-the-art literature review.
- ✓ Development of a new scheme exploiting two different approaches: a bio-inspired motion detection method and a BS algorithm based on color information.
- ✓ Review of objective segmentation evaluation methods in order to identify different metrics that could be used in this context [7].

2nd Year

- ✓ Exploration of the discrimination capability of different color spaces in the context of a GMM-based algorithm.
- ✓ Development of a robust and computationally efficient method, based on GMM [6], suitable for real-time applications.
- ✓ Development of a new approach to model the local texture at the pixel neighborhood.

3rd Year

- ✓ Proposal of a new method, based on a combination of local texture and pixel color representations, to address the problem of moving objects segmentation in night videos.

Next Year Planning

- **Writing of the PhD dissertation (Jul–Sep/2017)**
- **PhD defense (Oct–Nov/2017)**

Thesis Objectives

In this PhD research we want to further explore some of the most efficient approaches to propose a more robust algorithm.

Research Question:

Is it possible to improve widely used approaches, such as GMM, by proposing more robust algorithms while keeping complexity low?

Can this method be further improved in order to cope with more difficult cases, such as nighttime videos, even if sacrificing the complexity?

Publications Plan & International Conferences Presentations

- Martins, I., Carvalho, P., Corte-Real, L., Alba-Castro, J.L. (2016) ***Bio-inspired Boosting for Moving Objects Segmentation***. In: A. Campilho, F. Karray (eds) Image Analysis and Recognition. ICIAR 2016. Lecture Notes in Computer Science, vol. 9730, pp. 397-406, Springer, Cham.
  - **Oral presentation** at ICIAR 2016, July 13-15, 2016, Póvoa de Varzim, Portugal.
- Martins, I., Carvalho, P., Corte-Real, L., Alba-Castro, J.L. (2017) ***BMOG: Boosted Gaussian Mixture Model with Controlled Complexity***. In: L. A. Alexandre, J. S. Sánchez, J. M. F. Rodrigues (eds) Pattern Recognition and Image Analysis. IbPRIA 2017. Lecture Notes in Computer Science, vol. 10255, pp. 50-57, Springer Int. Publishing.
  - **Oral presentation** at IbPRIA 2017, June 20-23, 2017, Faro, Portugal.
- A paper to be submitted to an international journal is under preparation, with the provisional title ***COLBMOG: Texture-Based Segmentation for Night Videos***.

Results & Discussions

Bio-Inspired Boosting for Moving Objects Segmentation

1

(presented and discussed in 2016 Annual Defense)

The **main novelty** introduced is the fusion of **low-level information from the modeling of the human visual system** [5] with state-of-the-art methods used in **BS.**

BMOG: Boosted Gaussian Mixture Model with Controlled Complexity

2

The proposed solution explores **a novel classification mechanism** that combines:

- **Color space discrimination capabilities**
  - L*a*b* color space
  - Each channel component is analyzed independently and their decisions combined
- **Pixel classification with hysteresis**

and

- **Dynamic learning rate** for **background model update**.
  - adapted independently for each pixel
  - depends on the change of classification

Testing & Evaluation

- Tests with **CDnet 2014 dataset** [4]: **53 videos**, from **11 categories**, with **ground truth**.
- BMOG results submitted to the CDnet site to be evaluated and ranked.
- **BMOG compared** with **MOG2** [6], a widely used MoG based method, **RMoG**, a recent MoG based method and **SuBSENSE**, a top-rank state-of-the-art method.

Results show that BMOG consistently outperforms MOG2 method, and that it approaches top ranking, but much more complex, algorithms.

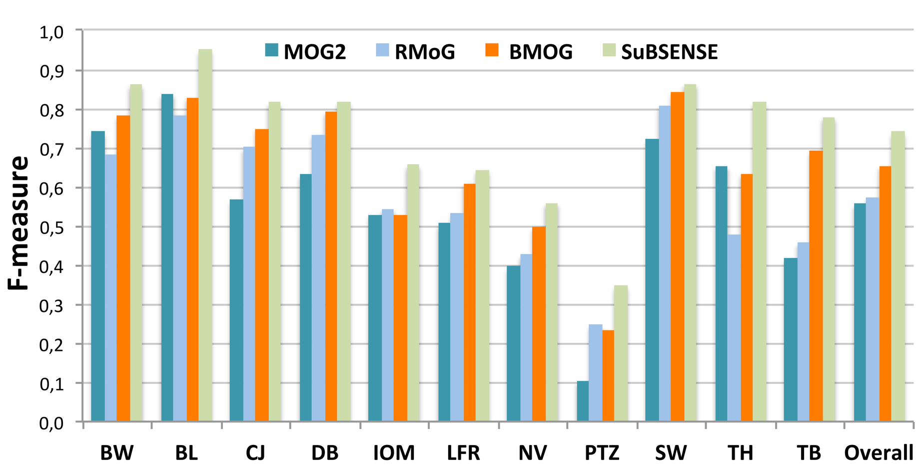
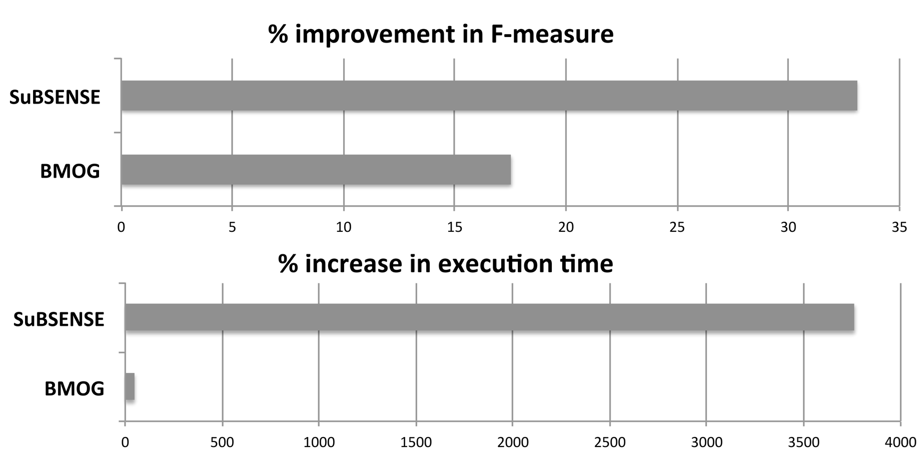


Fig. 1. Average F-measure for each category and across all categories.



- **Boosts the overall detection accuracy while keeping complexity low.**
- **Good choice for real-time applications.**

Fig. 2. Comparison of improvement in performance and increase in complexity for BMOG and a top-ranked method (reference: 0=MOG2).

COLBMOG: Texture-Based Segmentation for Night Videos

3

The proposed COLBMOG method is based on a **local texture feature** integrated with a **parametric background model** (BMOG).

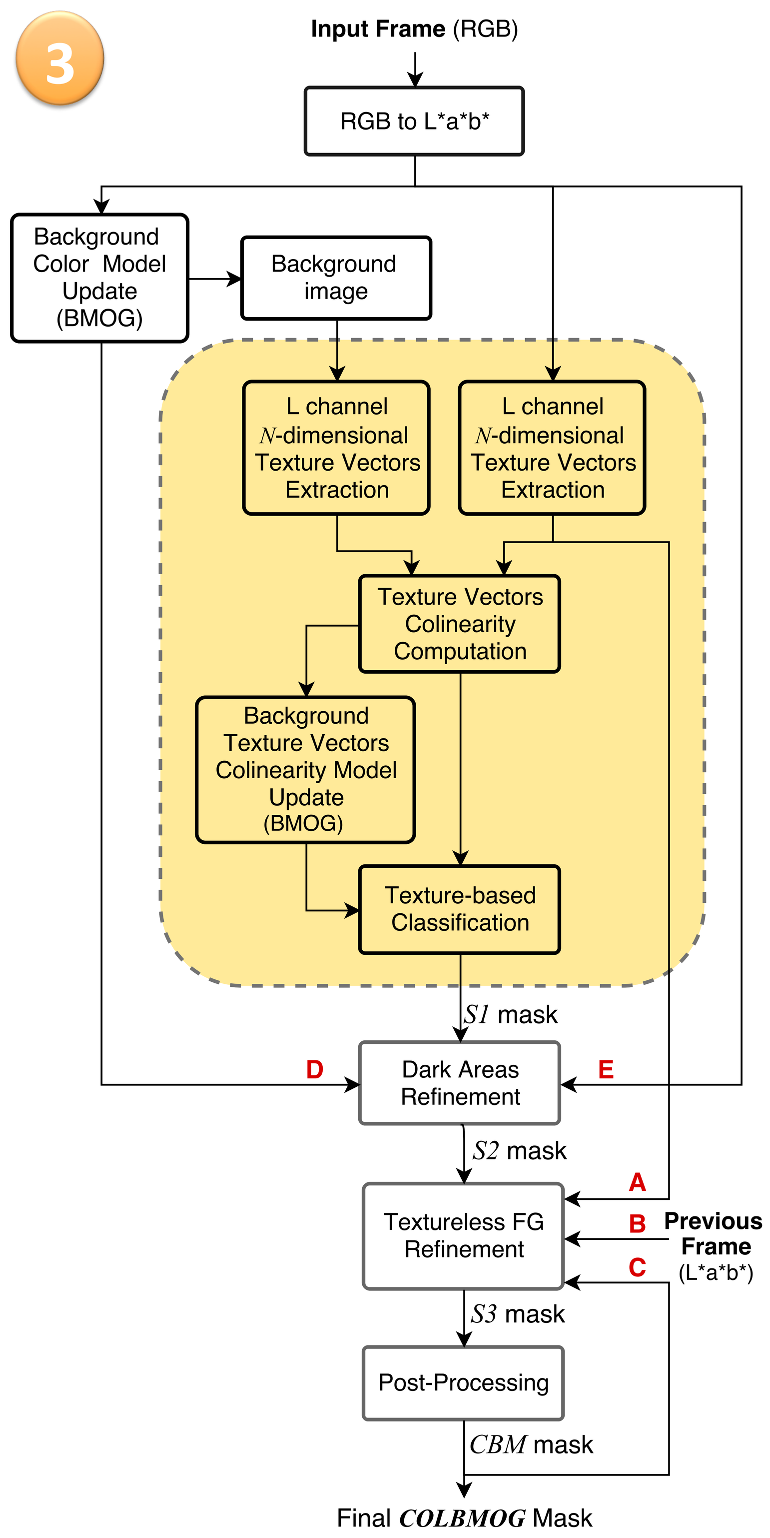


Fig. 4. COLBMOG block diagram.

Local texture feature:

***N*-dimensional vector** of pixel intensities extracted from *L* channel according to a pre-defined pattern

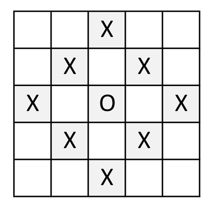


Fig. 3. Pattern of neighboring pixels representing local texture at pixel o.

Pixels whose associated texture vectors match the background image texture are classified as background (BG) and those that do not match are considered foreground (FG).

Measure of similarity between both textures at pixel j:

the **collinearity** between the corresponding texture vectors in the input frame, $\vec{x}_j$, and in the background image, $\vec{b}_j$.

Measure of collinearity: the **angle between the texture vectors**, $\theta(\vec{x}_j, \vec{b}_j)$.

A model of the collinearity between texture vectors of background pixels in successive frames based on BMOG is created and updated at every frame.

Testing & Evaluation: CDnet 2014 NightVideos (NV) category benchmark [4]

Table 1. F-measure across the overall set of videos for COLBMOG and EFIC, C-EFIC (the two top-ranked unsupervised methods in NV category) [4].

F-Measure	EFIC	C-EFIC	COLBMOG
Average	0,6548	0,6677	**0,7564**
St Dev	0,1245	0,1034	**0,0435**

RANKS FIRST in CDnet NV benchmark [4] **for the unsupervised methods.**

Not only the average F-measure is higher but also the standard deviation is significantly lower, meaning a more consistent performance across different challenges.

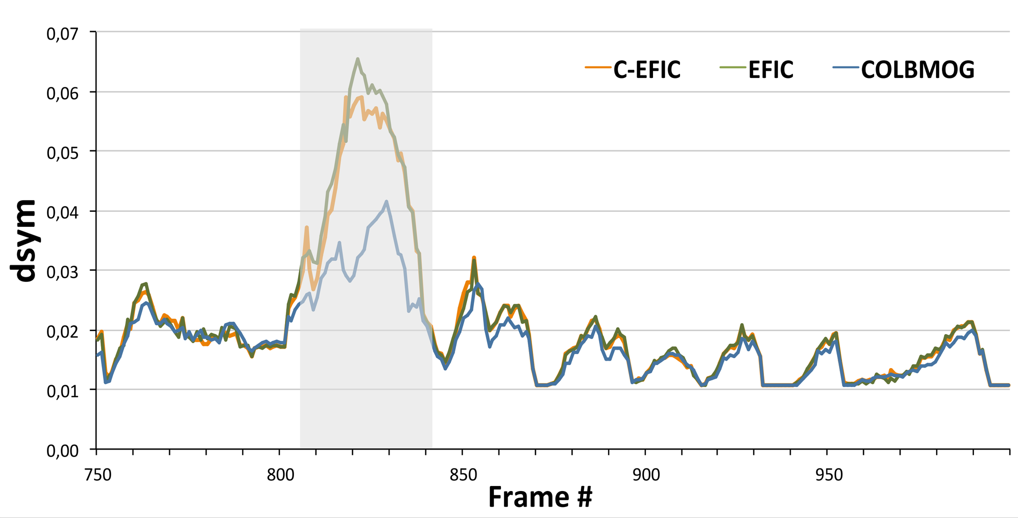


Fig. 5. Evolution of the error measure, **dsym**, for video *busyBoulvard*.

Fig. 5 and Fig. 6 show that all the algorithms tend to fail in the same frames, corresponding to the **most difficult situations**, and it is in these frames that **COLBMOG achieves a significant improvement**.

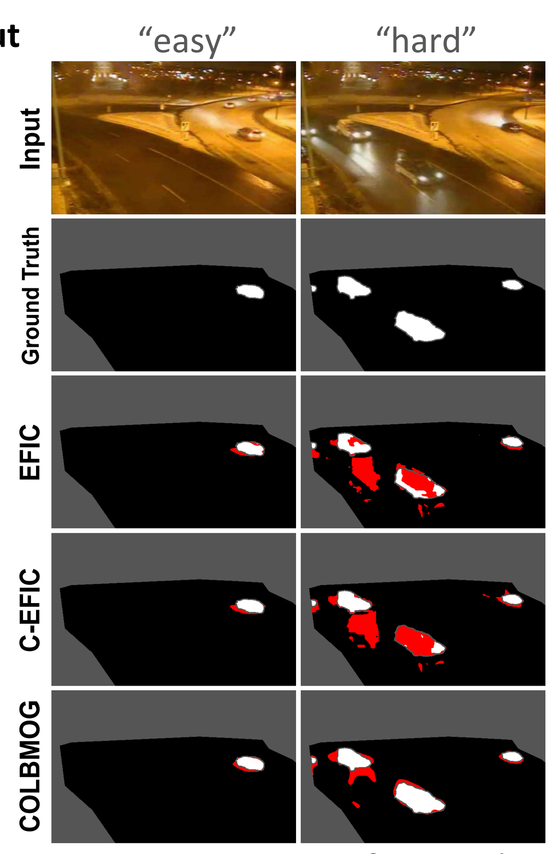


Fig. 6. Comparison of FG masks. Misclassified pixels are marked red.

References

[1] Bouwmans, T.: Traditional and recent approaches in background modeling for foreground detection: An overview. Computer Science Review 11, 31-66, May (2014)

[2] Stauffer, C., Grimson, E.: Adaptive background mixture models for real-time tracking. IEEE Int. Conf. Comput. Vision and Patt. Recogn. (CVPR) 2, 246–252, (1999)

[3] Wang, Y., Jodoin, P.-M., Porikli, F., Konrad, J., Benezeth, Y., Ishwar, P.: CDnet 2014: An Expanded Change Detection Benchmark Dataset. In Proc. CDW-2014, at CVPRW-2014, pp. 387–394 (2014)

[4] ChangeDetection.NET (CDNET), http://www.changedetection.net

[5] Benoit, A., Caplier, A., Durette, B., Herault, J.: Using human visual system modeling for bio-inspired low level image processing. Comput. Vis. Image Underst. 114,(7), 758–773 (2010)

[6] Zivkovic, Z., van der Heijden, F.: Efficient adaptive density estimation per image pixel for the task of background subtraction. Patt. Recogn. Lett 27(7), 773–780 (2006)

[7] Cardoso, J. S., Carvalho, P., Teixeira, L. F., Corte-Real, L.: Partition-distance methods for assessing spatial segmentations of images and videos. Comput. Vis. Image Underst. 113 (7), 811–823 (2009)