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, School of Telecommunications Engineering, U.Vigo, Spain

Luís Corte-Real, Associate Professor, Faculty of Engineering, U.Porto, Portugal

Motivation of the work

Moving objects detection and segmentation is a fundamental step in many vision systems. The detection and identification of changing or moving areas in a sequence of images is a fundamental pre-processing step for applications such as visual surveillance, smart environments, and video retrieval. Although subsequent processing may be different in each case, **moving objects segmentation is a critical factor for the success of the overall system**.

Developing robust and universal methods for unsupervised segmentation of moving objects in video sequences, also called **background subtraction (BS)**, has **proved to be a hard and challenging task**.

State-of-the-art methods [1] show good performance in a wide range of situations but **none has been able to fully deal with complex and challenging scenarios** that include poor lighting conditions, sudden illumination changes, shadows and parasitic background motion.

Some of the most widely used methods build a statistical model of background pixels, where the pixels are considered foreground if they do not fit in such model. Gaussian Mixture Models (GMM) are among the most popular [2].

Recent research has shown that methods appear to be complementary in nature, with the best-performing ones being beaten by combining several of them [3]. The combination of methods takes advantage of the different strengths of each individual method, resulting in algorithms that present a better performance but at the cost of significant increase in complexity and computational load.

The results reported by different authors have not been computed on a common dataset, making it hard or impossible to establish fair comparisons. Besides that, many datasets do not contain a balanced set of videos presenting real application challenges. Moreover, **metrics used to evaluate the average algorithms' performance do not reveal how they perform frame by frame**.

Thesis Objectives

In this PhD research we want to further explore some of the most efficient and widely used approaches (such as GMM) 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, even if sacrificing the complexity?

Methodology

The research question clearly involves the proposal of an algorithm with some expected complexity.

In this context, this research **follows the so-called Design Science approach** [4] to the development of the proposed Model (algorithm) that responds to the proposed research question. This model builds on the environment needs, researcher experience, literature and specific domain knowledge.

Validation builds on experimentation and on comparing obtained results with reference datasets widely used in the community. An exhaustive set of experiments has been conducted to evaluate the performance of the proposed methods compared with state-of-the-art methods.

A detailed analysis of the results, using different metrics, reveals how the proposed methods compare with state-of-the-art methods in different challenging scenarios.

The experiments for this research are conducted on the complete set of videos provided in the CDnet 2014 Dataset [5], depicting indoor and outdoor scenes captured in different scenarios and with different cameras, **containing a wide range of different challenges**. Testing and evaluation are performed using the ground truth segmentation provided along with the videos.

The **53 videos are grouped into 11 categories**, according to the type of challenge each represents: baseline, dynamic background, camera jitter, shadows, intermittent object motion, thermal, bad weather, low frame-rate, night, PTZ, air turbulence.

This **approach aims at producing repeatable experiments** that may contribute the the consistent development of the state-of-the-art in this domain.

Publications Plan

- As a result of the work on bio-inspired motion detection, a paper entitled **"Bio-Inspired Boosting for Moving Objects Segmentation" to be presented at ICIAR 2016** (Int. Conf. on Image Analysis and Recognition) to be held on July 13-15 (accepted for oral presentation).
- Presently, a paper is in preparation reporting the results of the improvement of the GMM algorithm, exploring the color space discrimination capability along with a different decision rule for the classification and a new dynamic learning rate to update the background model.
- A third paper is planned for the next months, proposing a new method combining the improved GMM algorithm with the texture based approach.

Next Year Planning

- We want to **explore the discrimination capability of different color spaces in the context of a GMM-based algorithm** [7]. We perfomed a statistically-based analysis, in different color spaces, of the correlation between decision items, such as total and individual component color-distances, to assess the discrimination capabilities of the different color spaces. We expect to propose a new decision rule for classification that allows improved performance with equivalent complexity.
- Some early experiments were made with a **new approach to model the local texture at the pixel neighborhood**. In the process, we looked at the values of a pixel and its N-1 neighbours as a vector in an N-dimensional space. These vectors are used as a feature representative of the texture. The similarity in texture at each pixel, is measured as the colinearity between: 1) the vector for the pixel in the current frame; 2) the vector for the corresponding pixel in the background. Exploratory tests revealed very promising results, namely robustness to variations in illumination and the presence of shadows.
- Finally, a **new method combining the two previous approaches is to be proposed**.

Results & Discussions

Bio-Inspired Boosting for Moving Objects Segmentation

We propose a new scheme, exploiting information obtained from two inherently different approaches: 1) a bio-inspired motion detection method, and 2) a BS algorithm based on pixel color information.

The main novelty introduced is that it is based on the fusion of low-level information from the modeling of the human visual system [6] with state-of-the-art methods used in BS.

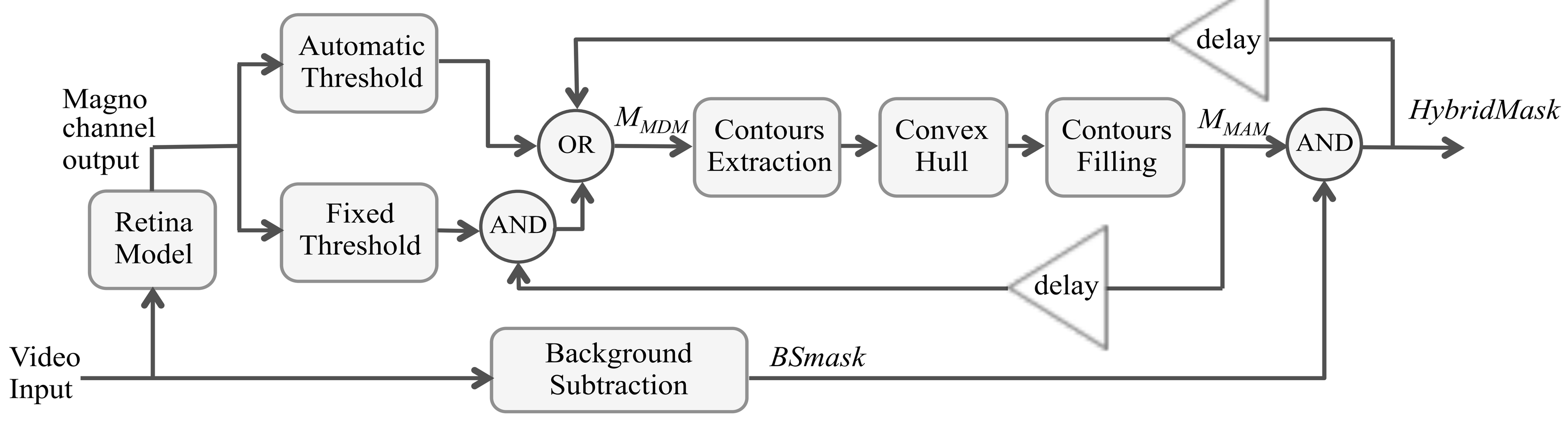

Fig. 2. Block diagram of the bio-inspired hybrid segmentation method

Experiments showed that the proposed method consistently improves the results, particularly in very complex situations, where the BS algorithms critically fail. This consistency is demonstrated below in the plots of d_{sym} and F-measure metrics.

Table 2. Average % of improvement in F-measure for each category and across all categories.

Category	MOG2	GMM	KNN	AMBER	Spectral 360	CwisarDH	SuBSENSE	FTSG
bad Weather	16.09	1.71	1.67	-0.11	-0.25	0.48	0.06	-0.04
baseline	11.23	1.10	0.79	0.50	-0.04	0.72	-0.11	-0.11
cameraJitter	20.87	0.64	0.73	0.04	-0.23	0.63	0.05	0.13
dynamicBackground	89.64	2.11	1.60	0.18	-0.20	1.20	0.04	-0.09
Interm.ObjectMotion	-2.60	0.20	-0.71	0.06	-0.41	0.55	-0.85	0.61
lowFrameRate	5.69	1.17	1.88	1.46	-0.01	2.81	-0.11	1.46
nightVideos	18.95	8.30	8.48	11.39	8.12	7.74	6.09	2.55
PTZ	17.49	2.76	2.98	1.56	3.45	4.04	0.28	1.69
shadow	9.43	1.70	2.47	0.63	0.27	0.74	-0.09	0.04
thermal	0.26	1.08	0.82	0.12	-0.32	0.75	0.23	0.32
turbulence	84.59	0.56	0.59	0.02	0.04	0.00	0.03	0.08
Overall	**19.79**	**1.76**	**1.76**	**0.97**	**0.64**	**1.40**	**0.39**	**0.45**

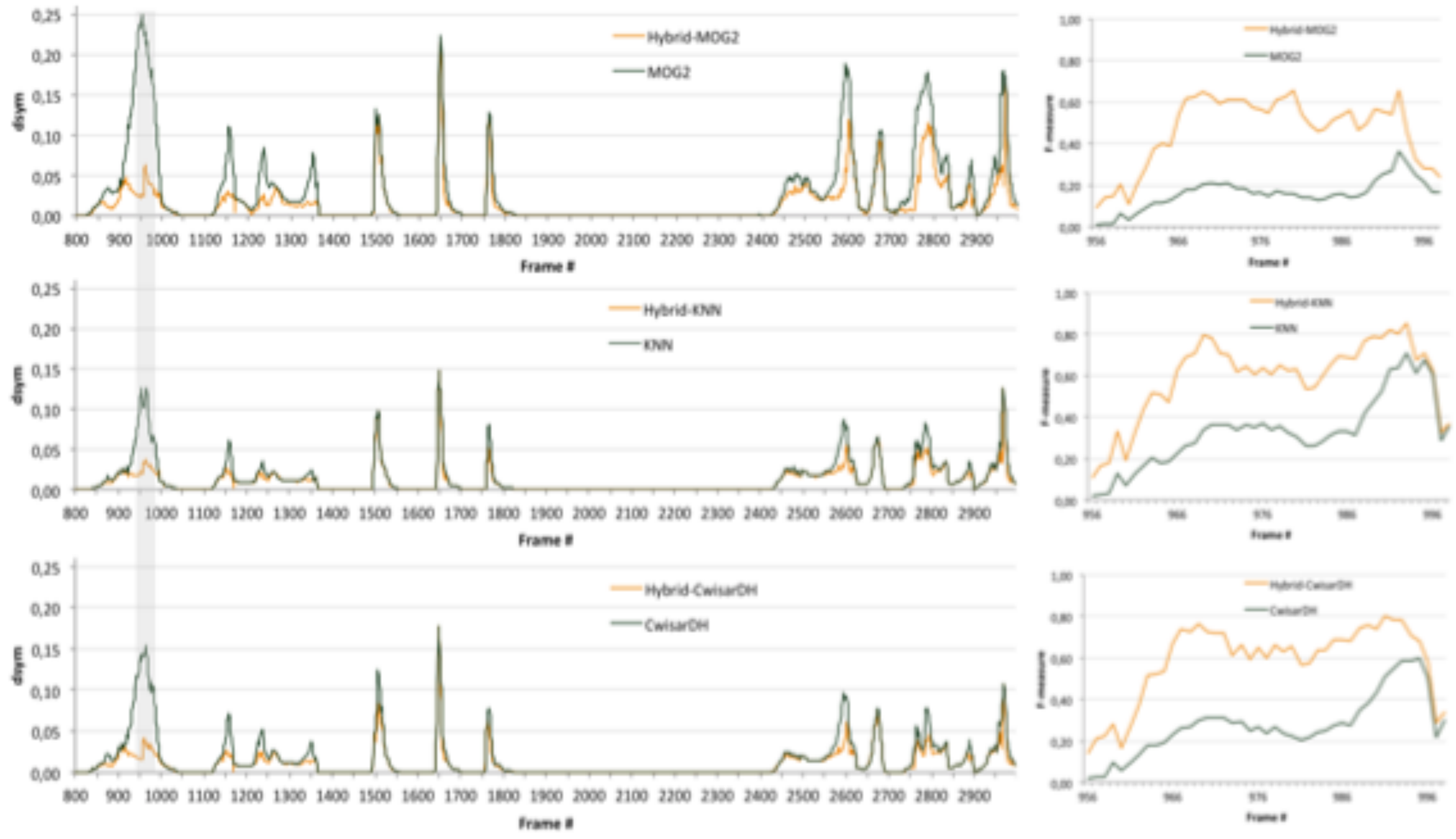

Fig. 3. Left: Evolution of d_{sym} from frame 800 to frame 2999. Right: Evolution of F-measure from frame 956 to frame 998 (shadowed region) of video *streetCornerAtNight*.

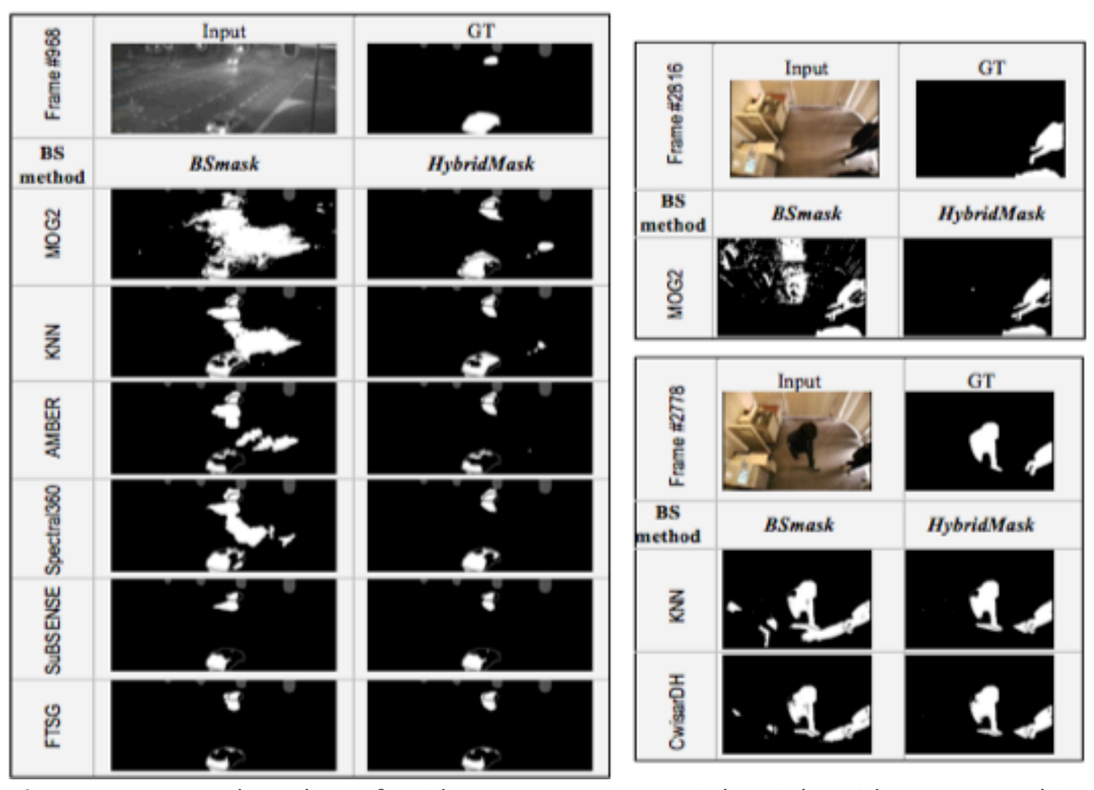

Fig. 5. Foreground masks. Left: video *streetCornerAtNight*. Right: video *copyMachine*.

References

[1] Bouwmans, T.: Traditional and recent approaches in background modeling for foreground detection: An overview. Computer Science Review, vol. 11, pp. 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), vol. 2, pp. 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] Hevner, A. R., March, S. T., Park, J. & Ram, S.. Design science in information systems research. MIS Q., vol. 28, pp. 75-105, (2004)

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

[6] Benoit, A., Caplier, A., Durette, B., Herault, J.: Using human visual system modeling for bio-inspired low level image processing. Comput. Vis. Image Underst., vol. 114, no. 7, pp. 758–773 (2010)

[7] Zivkovic, Z., van der Heijden, F.: Efficient adaptive density estimation per image pixel for the task of background subtraction. Patt. Recogn. Lett., vol. 27, no. 7, pp. 773–780 (2006)