Context-supported Road Information for Background Modeling

Marcelo Mendonça, Luciano Oliveira
Federal University of Bahia (UFBA)
Salvador, Brazil
www.ivisionlab.eng.ufba.br

Fig. 1. Background modeling and the intersection problem: first row shows subsequent moments of a traffic video obtained from the monitoring of an intersection. Because the traffic light is switching between each opposite direction, there are always occlusion due to vehicles stopped in a red light. Second row shows the correspondent background models produced by the proposed method, CRON. The method is capable of learning the background whenever background pixels become visible, so that background consistency is preserved even when the vehicles stop again.

***Abstract*—Background subtraction methods commonly suffers from incompleteness and instability over many situations. If one treats fast updating when objects run fast, it is not reliable to modeling the background while objects stop in the scene, as well; it is easy to find examples where the contrary is also true. In this paper we propose a novel method – designated Context-supported ROad iNformation (CRON) for unsupervised background modeling, which deals with stationary foreground objects, while presenting a fast background updating. Differently from general-purpose methods, our method was specially conceived for traffic analysis, being stable in several challenging circumstances in urban scenarios. To assess the performance of the method, a thorough analysis was accomplished, comparing the proposed method with many others, demonstrating promising results in our favor.**

***Keywords*-Background modeling; traffic analysis; surveillance videos.**

I. Introduction

By processing videos provided by a surveillance camera, a variety of traffic activities can be automated and improved. In this context, several methods have been proposed to automatically perform traffic flow measurement [1], [2], vehicle classification [1], [3], traffic-violation detection [4], intelligent traffic light control [5], road area segmentation [6] and car parking monitoring [7], just to cite a few. An important issue is that, despite the difference in the final purposes, many of these solutions are based on background subtraction (BGS). Broadly speaking, BGS consists in separating the permanent components of the environment (e.g. ground, buildings, trees) from the objects that eventually appear in the scene (e.g. cars, people, animals). In the context of traffic flow analysis, vehicles compose the foreground while the background is comprised of the road, most of the time.

There are several BGS methods, which can be categorized by their approaches: basic methods, based on direct inter-frame difference or simple statistical inferences (e.g., mean, median or variance) [8], [9]; statistical methods, based on Gaussian distributions [10], [11]; fuzzy-based methods [12], [13], which are based on color and texture features [14]; non-parametric methods [15], based on eigenvalue analysis [16]; and neural [17] and neural-fuzzy [18] methods (Refer to [19] for a survey). The difference among them resides mainly in a trade-off between precision on foreground detection and computational load. Nevertheless, a drawback is common to all of them: the background modeling does not yield a reliable model, specially as dealing with eventually stationary foreground, i.e., objects that stop in the scene for a while but does not belong to the background itself (e.g., cars approaching to an intersection). The method proposed in [11], for example, is fast enough to update the background when a previous static object starts moving, but, in contrast, the foreground is quickly incorporated by the background if the contrary occurs.

There are too few works which specifically address stationary foreground problems in background modeling for traffic analysis. In [7], the authors propose an algorithm to detect illegally parked vehicles in traffic video. The proposed method requires a previously marked region of interest (ROI), from which an initial background model is obtained, calculating the median of the gray-level intensity of the pixels inside the ROI. The running average of the pixel intensities is calculated to update the background along the frames, and the absolute difference between the initial and the updated background yields the foreground. Tracking is then performed in order to detect if the foreground becomes static. This method has the shortcoming of massively relying on tracking, which is difficult to benefit in presence of crowded scenes. In [20], Hu *et al.* propose a background updating scheme claimed to be robust to three mainly problems: sudden camera perturbation, illumination changes and stationary foreground. To tackle those problems, some techniques are used, such as: Euclidean transformation combined with optical flow to stabilize the images, normalized histograms to reduce effects of luminance, and a background modeling based on pixel-value zone distributions. The background is taken from the pixels with the most frequent gray-level intensities in the distribution. However, if a foreground object remains static most of the time, the method fatally fails. Bhandarkar and Luo [21] propose to model the background as a temporal moving median of each pixel. In a first stage, a set of rules and operations are applied pixel-wise to verify if a pixel belongs or not to the background. To avoid misclassification of motionless foreground objects, a second stage involving speed estimation, position prediction and inter-object correspondence analysis is also performed. According to the authors, the proposed system does not present reliable results where previous static objects or heavy traffic conditions are involved; yet, like the other methods, that latter one only produces a gray-level background model, and valuable color information is thus discarded.

Contributions: In this paper we present an unsupervised method called Context-supported ROad iNformation (CRON), specially designed for background modeling in traffic scenes. Our method innovates by providing a fast background updating along with an adaptive scheme for road color learning. These characteristics ensure the robustness of the method with respect to foreground objects that eventually stop in the scene, a situation that usually prevent common BGS methods to produce consistent background models (see Fig. 1 to better understand this issue). Our novel approach is reached by means of the following strategy: Firstly, BGS based on approximated median is performed to extract a foreground mask, as well as to generate an earlier color background model. Next, a simple and fast vehicle detection technique called *vehicle filtering* is performed on the foreground mask in order to filter out those blobs most likely to correspond to vehicles in the scene. Logically, by detecting vehicles one can also detect parts of the road. The revealed road parts provide information about the road color, which our method uses adaptively to update the background model by maximizing the road color in the image – the predominant color in the background of a traffic image. The system is outlined in Fig. 2.

II. Background modeling steps

A. Obtaining an earlier background model

The rationale of CRON resides in using road color information as reference to adaptively model the background, in order to better distinguish among background and foreground objects, when dealing with traffic videos. For that aim, an initial requirement is to extract some road information, as well as to obtain a starting background model which will be posteriorly updated. To cope with this issue, our method is initially based on the approximated median BGS technique, proposed in [8], to separate the foreground from the background at the beginning of the processing (by now, without concerning for the background consistency problem). The approximated median algorithm is an improvement of the original median based BGS, which addresses the memory requirement problem. In the approximated form, instead of storing each pixel value along the frames to calculate the median, the background is obtained by increasing or decreasing pixel intensities, depending on whether its value becomes higher or lower from the current frame to the next. This is given by

$$BG_k = \begin{cases} BG_{(k-1)} + u, & \text{if } Frame_k - BG_{(k-1)} > 0 \\ BG_{(k-1)} - u, & \text{otherwise} \end{cases} \tag{1}$$

where BG denotes pixels in background, u is the background updating factor, and $k = 1, 2, ..., N$ is the number of frames to be analyzed.

Additionally, a foreground mask can be obtained by

$$FG_k = \begin{cases} 1, & \text{if } \left|Frame_k - BG_{(k-1)}\right| > \rho \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

where FG denote pixels in the foreground, and ρ is the threshold to determine the foreground.

The motivation to use the approximated median BGS as support for our background modeling method (see Fig. 2) is mainly because it attends the on-the-fly premises of our goals. Besides, when in presence of relatively fast moving objects, by adequately choosing u and ρ, the approximated median yields a background model good enough for our purposes at this

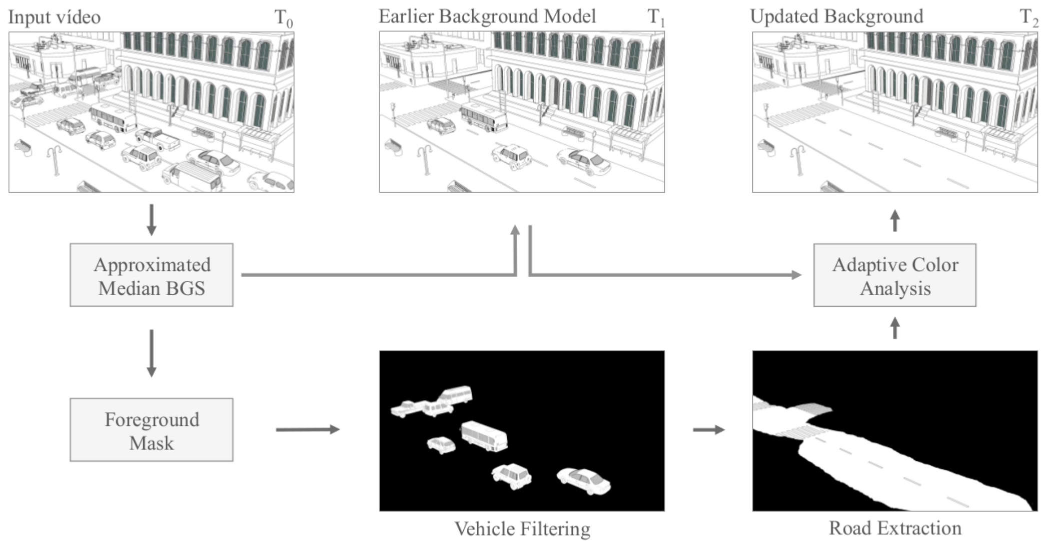


Fig. 2. Outline of the method. CRON is comprised of the following steps: an approximated median background subtraction method which earlier separates the input video (in time T_0) in foreground and background (that one in time T_1); next, the foreground is analyzed with the aim at finding vehicles, which are further used to detect road parts; finally, an adaptive process of color analysis extracts road information in order to produce a new background model (in time T_2) by updating the earlier one (in time T_1). Note that T_n is time and not frame.

point. Indeed the quality of the background model produced at this point is not critical, since it will be posteriorly updated by CRON, as will be described in the next sections.

B. Vehicle filtering

Methods in the same category of CRON, which provide background modeling, usually do not deal properly with foreground objects when they become slow or static in the video sequence. This mainly happens due to the generalist approaches of these methods, focused on just temporal aspects of the pixel behaviour. In other words, since such methods do not use any contextual information in order to figure out how the foreground and background pixels are, they are not capable of distinguishing among them if these two kinds of pixels behave similar for a while. To address that question, CRON adopts a context-aware approach. For that, the proposed method starts by analyzing the foreground mask provided by Eq. 2 to filter those objects more likely to be cars. For that, each blob in the mask (black regions) is distinguished by following a chain of criteria (see Fig. 3):

- **Solidity** consists in a measure of how much pixels in the blob are also pixels of the convex hull which envelops that blob. The idea is that cars tend to be convex objects and, therefore, if a given object presents low solidity, it is probably not a car. This first criterion mainly serves to eliminate those blobs derived from groups of people that are walking near to each other. However, the solidity alone is not enough to determine if an object is a car. For example, people walking alone with arms near to the body appear to be reasonably convex (as seen with the person inside the green hull in Figure 3). Then, to deal with such situations, the algorithm analyzes the blob under the next criterion.
- **Aspect ratio**. In traffic surveillance videos, foreground is normally comprised of vehicles passing on the roads, as well as people crossing these roads or walking on the sidewalks. In this context, a discriminant feature to distinguish cars from isolated people is the aspect ratio, i.e., the relation between height and width of an object. Since the height of a common person is not smaller than three times the width, in situations where it is not true the object is probably a car (green arrows in Fig. 3). This rule, however, is not valid for long vehicles, like bus or trucks, or even small cars queued (yellow arrows in Fig. 3). For those cases, the blobs are further examined under the next criterion.
- **Orientation**. Taking the x-axis as reference (white segments in Fig. 3), the angle formed by the longest axis of the blobs is somewhat different when it comes from people (red arrow in Fig. 3) or vehicles. While the former presents nearly 90 degrees, the latter follows the orientation of the roads, which are almost never vertical from the common viewpoints of surveillance cameras. Thus, by choosing the correct angles, the orientation criterion can help us to further identify vehicles in the foreground mask.

After passing through this triage process, the remaining objects – taken as being vehicles – are used to detect road regions. For that, only the pixels at the bottom of the vehicle convex hulls (blue regions in Fig. 3) are selected. The goal is to avoid misclassification, since the inferior parts of the vehicles are always overlapping image road regions, while it is not possible to happen for the upper parts. At the end of the vehicle filtering, the selected pixels are then accumulated along the frames. Road color information extracted from these pixels is the basis for the background update, a process that is carried out by means of an adaptive analysis, as will be described in the next section.

C. Background update based on road color adaptive analysis

In order to analyze the color aspect of the road pixels, a descriptor called gray-amount (Algorithm 1, Line 9) was conceived. This descriptor was proposed to rule the color part of CRON, and it consists in the mean of two measures:

- **Pixel grayness**. This measure is reached by calculating, for each pixel, the average of the RGB inter-channel square differences (Line 7).
- **Color pixel distance to the road color**. This measure consists in the euclidean distance from the pixel intensity to the road color (Line 8).

The first measure aims to favour gray pixels, since they have similar values in each channel. Conversely, the greater the inter-channel difference, the lower this measure. The rationale for that is because road pavement usually presents gray color. With respect to the second measure, the pixel intensity is computed by taking the mean of its RGB values; next, it is calculated the distance to the road intensity. In CRON, road color is adaptively obtained by extracting samples of road pixels from the earlier background model. This is possible due to the road portions revealed at the filtering process presented in Sec. II-B. In this case, we calculate the weighted average of the road pixels sampled in the earlier stage, where the weights are given by the correspondent accumulation of the detected vehicles along the frames. The idea is to give more weight to the colors of the road regions where the heavier traffic occurs. Another aspect of this descriptor is that we have raised the RGB inter-channel differences to square power. The goal is to further penalize the pixels that are not gray, but whose the average of the channels is close to the average intensity of the road pixels. Algorithm 1 delineates each step of the proposed background modeling with this adaptive color process.

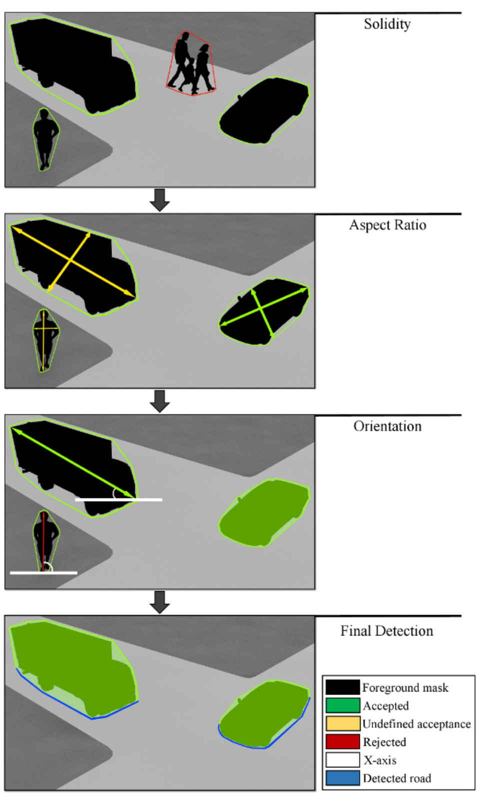


Fig. 3. Road detection process. From top to bottom: foreground blobs are evaluated by the solidity criterion, which measures the area of intersection among the blob and its convex-hull; the aspect ratio of this convex-hull is analyzed: if an object meets this criterion, it is considered a car; the relative orientation to the x-axis is further considered in order to distinguish vehicles from people; at the end, the road regions are taken as the bottom of the convex-hulls of the objects passed in the criterion chain.

Algorithm 1 CRON: complete algorithm

Input: Traffic video
Output: Background model $newBG$

1: */* Main Routine */*
2: **for** k = 1 to K frames **do**
3: $[BG_{(k)},\ FG_{(k)}] \leftarrow$ *approximatedMedianBGS(k)*
4: $Rmap_{(k)} \leftarrow$ *vehicleFiltering(*$FG_{(k)}$*)*
5: $roadColor \leftarrow$ *weightedMean(*$Rmap_{(k)}, BG_{(k)}$*)*
6: */* Pixelwise operations */*
7: $G'_{(k)} \leftarrow \dfrac{\displaystyle\sum_{\substack{i=1,\\ C_i \in \{R,G,B\}}}^{2} \sum_{\substack{j=i+1,\\ C_j \in \{R,G,B\}}}^{2} \left(BG_{C_i} - BG_{C_j}\right)_k^2}{3}$
8: $G''_{(k)} \leftarrow \left| \frac{BG_{R(k)} + BG_{G(k)} + BG_{B(k)}}{3} - roadColor \right|$
9: $G_{(k)} \leftarrow \frac{G'_{(k)} + G''_{(k)}}{2}$ ▷ Gray Amount
10: */* Update */*
11: **if** $(G_{(k)} > G_{(k-1)})$ **then**
12: $newBG_{(k)} \leftarrow BG_{(k)}$
13: **end if**
14: **end for**

At the end, as can be seen from Line 10, CRON controls the updating of the earlier background model by maximizing the gray amount for each of their pixels. This way, if a previous static foreground object starts moving, then the gray amount in that region increases due to the road pixels being revealed, which results in a fast background updating. On the other hand, if a moving object suddenly stops in somewhere in the image, the gray amount in that region gradually decreases, causing the immediately interruption of the updating process there so that the background consistency is preserved.

III. Experimental analysis

A. Vehicle detection performance

As mentioned, the vehicle detection heuristics addressed in Section II-B is expected to have low computational cost to attend the system *on-the-fly* requirements. Besides, the method

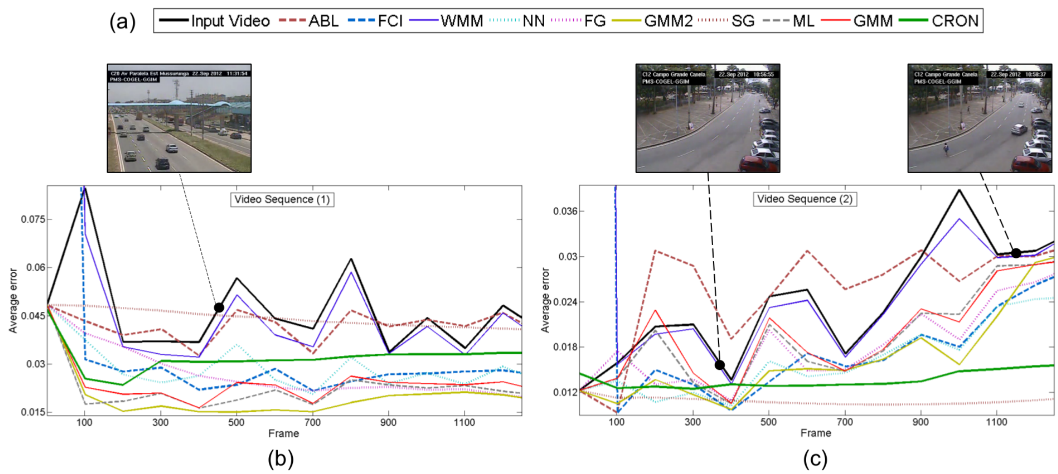


Fig. 4. Average error in background modeling - Part 1. In (a), the legend for the graph curves; (b) and (c) show the curves related to, respectively, video sequences 1 and 2. Black curve represents the raw video, and the remaining ones are the compared methods.

is not intended to present a vehicle detection performance comparable to state-of-the-art methods, since the goal is simply to detect some cars in order to reveal samples of road pixels. The most important, in this case, is to reach an as low as possible false positive rate, since false positives mean noise in the road color extraction.

To assess the performance of the method in vehicle detection, we have gathered 29 traffic videos including different scenarios (e.g. highways, intersections) and situations (e.g. free traffic, traffic jam with vehicles and sidewalk people). Although it is of our knowledge that there exist some public data sets for surveillance traffic available in the Internet, the gathered data set provides us with unique characteristics of difficulty, object occlusions, road with variable shapes, shadows and lighting variation.

The performance was measured thus by the matching of the detected pixels of the proposed method with the ground truth annotated for the road regions in the video data sets. The method obtained an average precision of 84%, an average recall of 84% and an average accuracy of 66% (considering a threshold of 0.005 for the accumulated value).

B. Background modelling performance

An experimental analysis is presented, comparing CRON with 9 other methods[1]. They are: adaptive background learning (ABL) [22], fuzzy Choquet integral (FCI) [13], weighted moving mean (WMM), self organization through artificial neural networks (NN) [17], fuzzy Gaussian (FG) [23], two versions of Gaussian mixture model — (GMM) [11] and (GMM2) [19], simple Gaussian (SG) [10], and multi-layer (ML) [14].

To assess the performance of the methods in producing reliable background models, we have gathered four traffic videos, each video containing a particular situation that imposes different difficulties to the methods. Table I summarizes the characteristics of the selected videos.

TABLE I
CHARACTERISTICS OF THE VIDEOS FOR EVALUATION.

Video	Characteristic
#1	Vehicles run uninterruptedly with normal traffic conditions.
#2	An abnormal situation occurs from frame 700, when a car slows down until stopping, parks on the roadside, and then reverses looking for a parking spot.
#3	An intersection where cars stop in one of the roads, while cars on the other road are running; from frame 445, the traffic light switches and the situation inverts.
#4	A moderate traffic jam occurs between frames #1 and #250, forcing some vehicles to stop, while others remain moving forward; at frame 500, all vehicles stop for a pedestrian crossing.

The comparative analysis was made as follows: i) First, a background model to be used as reference was manually extracted from each video sequence; ii) after that, binary masks were generated delimiting the road regions for each background reference; iii) next, for each input video frame, a background model was produced by each BGS method to be evaluated; iv) at the end, the generated models were then compared with the corresponding background reference, considering only the road regions delimited by the road masks, for each frame. For each frame, the average error is taken as the (pixel-wise) mean of the difference between the reference background and the generated background models. The results obtained from this analysis are depicted in Figures 4 and 5, where the average error presented for each method has been plotted. To better view the plots, the initial parts of the curves are not completely shown, since it corresponds to the stabilization phase, where too high errors usually occur.

[1]Implementations available at code.google.com/p/bgslibrary/.

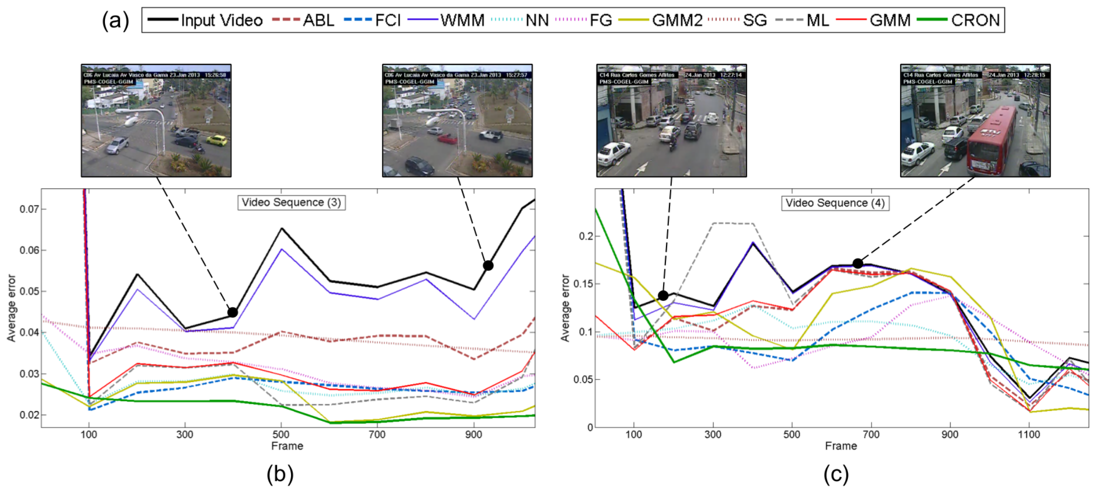


Fig. 5. Average error in background modeling - Part 2. In (a), the legend for the graph curves; (b) and (c) show the curves related to, respectively, video sequences 1 and 2. Black curve represents the raw video, and the remaining ones are the compared methods.

C. Discussion

Observing the plots of video #1 (Fig. 4.b), most of the methods performs slightly better than CRON. This was expected since in the cases where the foreground is always moving fast, those methods based only on temporal analysis can easily distinguish foreground from background pixels (FCI, NN, FG, GMM, GMM2, ML). CRON, on the other hand, accumulates some small errors due to foreground pixels that eventually present a gray amount higher than the background color. Despite that, the performance of CRON is better than WMM, ABL (which appear to just follow the input) and SG, which presents a very slow background updating.

In video #2 (Fig. 4.c), when the abnormal situation occurs, almost all methods are prone to incorporating motionless foreground pixels as background ones. Only SG and CRON yield stable background models in this situation. However the former reaches this result by means of a poor background update, as happens in video #1.

From video #3, in Fig. 5.b, it is noteworthy that the other methods suffers from the intersection problem, when the cars on a road start moving and the others become static: leading these methods to always misclassifying foreground as background. After the stabilization phase, CRON's performance is always the best.

The plots corresponding to the video #4 in Fig. 5.c show that some methods, such as FCI and FG, are sometimes capable of dealing with motionless foreground objects for a short time interval, as is the case in the initial part of the video. However, if the foreground becomes static for a longer period, as when the vehicles stop on the crosswalk, in the middle part of the video, then these methods also fail in the background modeling. CRON, on the other hand, has demonstrated to be able to produce reliable background models along all situations. The stability presented in this modeling process is an interesting advantage, specially for those applications aimed to segment background objects, as road detection methods.

The characteristics of CRON are further evidenced by the summarized data shown in Fig. 6, where CRON's mean and standard deviation of the error remain low throughout all the frames, demonstrating the stability of the method. Moreover Fig. 7 compares the (manually extracted) background references with the models generated by the best four competing methods (GMM2, SG, ML and FCI) and CRON. It is noteworthy how CRON's models are always robust to motionless foreground objects.

IV. Conclusions

In this paper, a novel unsupervised method of background modeling based on context awareness was proposed. The method, so called CRON, is motivated by an adaptive learning of road color, taking into consideration the gray amount of the road and its difference from the vehicles. Blobs of the vehicles were defined by three criteria, involving solidity, aspect ratio and orientation with respect to x-axis. The proposed method demonstrated stability in difficult situations requiring fast updating of the background and detection of stationary objects on the road. We are working now on taking into consideration local adaptiveness of color regions in order to cope with high variations of the road color (e.g, as in shadows).

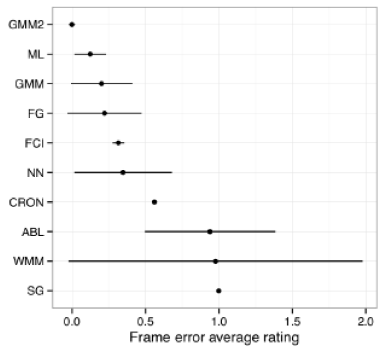
Acknowledgment

The authors would like to thank this colleague and this financing institute.

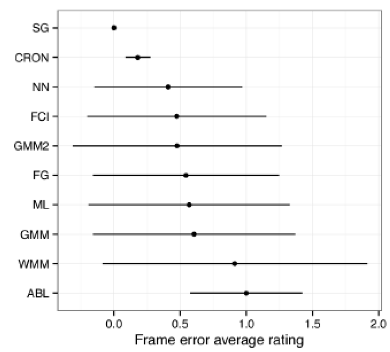
References

[1] L. Unzueta, M. Nieto, A. Cortes, J. Barandiaran, O. Otaegui, and P. Sanchez, "Adaptive multicue background subtraction for robust vehicle counting and classification," *Intelligent Transportation Systems, IEEE Transactions on*, vol. 13, no. 2, pp. 527–540, June 2012.

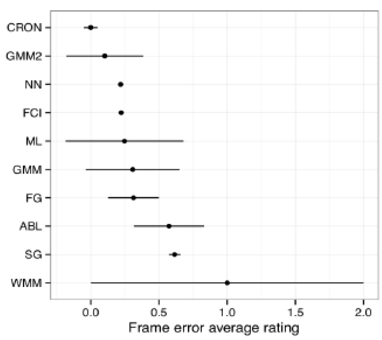
[2] A. Sobral, L. Oliveira, L. Schnitman, and F. D. Souza, "Highway traffic congestion classification using holistic properties," in *Signal Processing, Pattern Recognition and Applications (SPPRA), 10th IASTED International Conference on*, 2013.



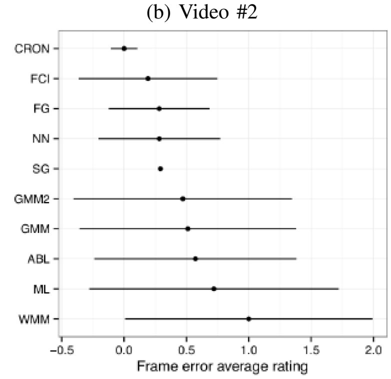
(a) Video #1



(b) Video #2



(c) Video #3



(d) Video #4

Fig. 6. Mean and standard deviation of the error presented in Figs. 4 and 5.

[3] Z. Chen, N. Pears, M. Freeman, and J. Austin, "Road vehicle classification using support vector machines," in *Intelligent Computing and Intelligent Systems, 2009. ICIS 2009. IEEE International Conference on*, vol. 4, Nov 2009, pp. 214–218.

[4] J. Vijverberg, N. de Koning, J. Han, P. de With, and D. Cornelissen, "High-level traffic-violation detection for embedded traffic analysis," in *Acoustics, Speech and Signal Processing (ICASSP), IEEE International Conference on*, vol. 2, April 2007, pp. II–793–II–796.

[5] M. Malhi, M. Aslam, F. Saeed, O. Javed, and M. Fraz, "Vision based intelligent traffic management system," in *Frontiers of Information Technology (FIT), 2011*, Dec 2011, pp. 137–141.

[6] M. Santos, M. Linder, L. Schnitman, U. Nunes, and L. Oliveira, "Learning to segment roads for traffic analysis in urban images," in *Intelligent Vehicles (IV), IEEE Symposium on*, 2013, pp. 527–532.

[7] H. Smitha and V. Palanisamy, "Detection of stationary foreground objects in region of interest from traffic video sequences," *International Journal of Computer Science Issues (IJCSI)*, vol. 9, p. 194, 2012.

[8] N. McFarlane and C. Schofield, "Segmentation and tracking of piglets in images," *Machine Vision and Applications*, vol. 8, no. 3, pp. 187–193, 1995. [Online]. Available: http://dx.doi.org/10.1007/BF01215814

[9] R. Cucchiara, C. Grana, M. Piccardi, and A. Prati, "Detecting moving objects, ghosts, and shadows in video streams," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 25, no. 10, pp. 1337–1342, Oct 2003.

[10] Y. Benezeth, P.-M. Jodoin, B. Emile, H. Laurent, and C. Rosenberger, "Review and evaluation of commonly-implemented background subtraction algorithms," in *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on*, Dec 2008, pp. 1–4.

[11] Z. Zivkovic, "Improved adaptive gaussian mixture model for background subtraction," in *Pattern Recognition (ICPR), Proceedings of the 17th International Conference on*, vol. 2, 2004, pp. 28–31 Vol.2.

[12] H. zhang and D. xu, "Fusing color and texture features for background model," in *Fuzzy Systems and Knowledge Discovery*, ser. Lecture Notes in Computer Science, L. Wang, L. Jiao, G. Shi, X. Li, and J. Liu, Eds. Springer Berlin Heidelberg, 2006, vol. 4223, pp. 887–893. [Online]. Available: http://dx.doi.org/10.1007/11881599_110

[13] F. El Baf, T. Bouwmans, and B. Vachon, "Fuzzy integral for moving object detection," in *Fuzzy Systems (FUZZ-IEEE), IEEE International Conference on*, 2008, pp. 1729–1736.

[14] J. Yao and J. Odobez, "Multi-layer background subtraction based on color and texture," in *Computer Vision and Pattern Recognition (CVPR), IEEE Conference on*, 2007, pp. 1–8.

[15] Y. Goyat, T. Chateau, L. Malaterre, and L. Trassoudaine, "Vehicle trajectories evaluation by static video sensors," in *Intelligent Transportation Systems Conference, 2006. ITSC '06. IEEE*, Sept 2006, pp. 864–869.

[16] Z. Xu, P. Shi, and I. Y.-H. Gu, "An eigenbackground subtraction method using recursive error compensation," in *Proceedings of the 7th Pacific Rim Conference on Advances in Multimedia Information Processing*, ser. PCM'06. Berlin, Heidelberg: Springer-Verlag, 2006, pp. 779–787. [Online]. Available: http://dx.doi.org/10.1007/11922162_89

[17] L. Maddalena and A. Petrosino, "A self-organizing approach to background subtraction for visual surveillance applications," *Image Processing, IEEE Transactions on*, vol. 17, no. 7, pp. 1168–1177, 2008.

[18] ——, "A fuzzy spatial coherence-based approach to background/foreground separation for moving object detection," *Neural Computing and Applications*, vol. 19, no. 2, pp. 179–186, 2010. [Online]. Available: http://dx.doi.org/10.1007/s00521-009-0285-8

[19] T. Bouwmans, F. El Baf, and B. Vachon, *Statistical Background Modeling for Foreground Detection: A Survey*. Handbook of Pattern Recognition and Computer Vision, 2010, vol. 4, ch. 3, pp. 181–199.

[20] C. Hu, D. Li, and J. Liu, "Automatic background updating for video-based vehicle detection," in *International Symposium on Photoelectronic Detection and Imaging*, vol. 6623, 2007. [Online]. Available: http://dx.doi.org/10.1117/12.791497

[21] S. Bhandarkar and X. Luo, "An efficient background updating scheme for real-time traffic monitoring," in *Intelligent Transportation Systems, 2004. Proceedings. The 7th International IEEE Conference on*, Oct 2004, pp. 859–864.

[22] C. Zhang, S. ching Chen, M. ling Shyu, and S. Peeta, "Adaptive background learning for vehicle detection and spatio-temporal tracking," in *IEEE Pacific-Rim Conf. on Multimedia (PCM)*, 2003, pp. 15–18.

[23] M. H. Sigari, N. Mozayani, and H. R. Pourreza, "Fuzzy running average and fuzzy background subtraction: concepts and application," *International Journal of Computer Science and Network Security*, vol. 8, no. 2, pp. 138–143, 2008.

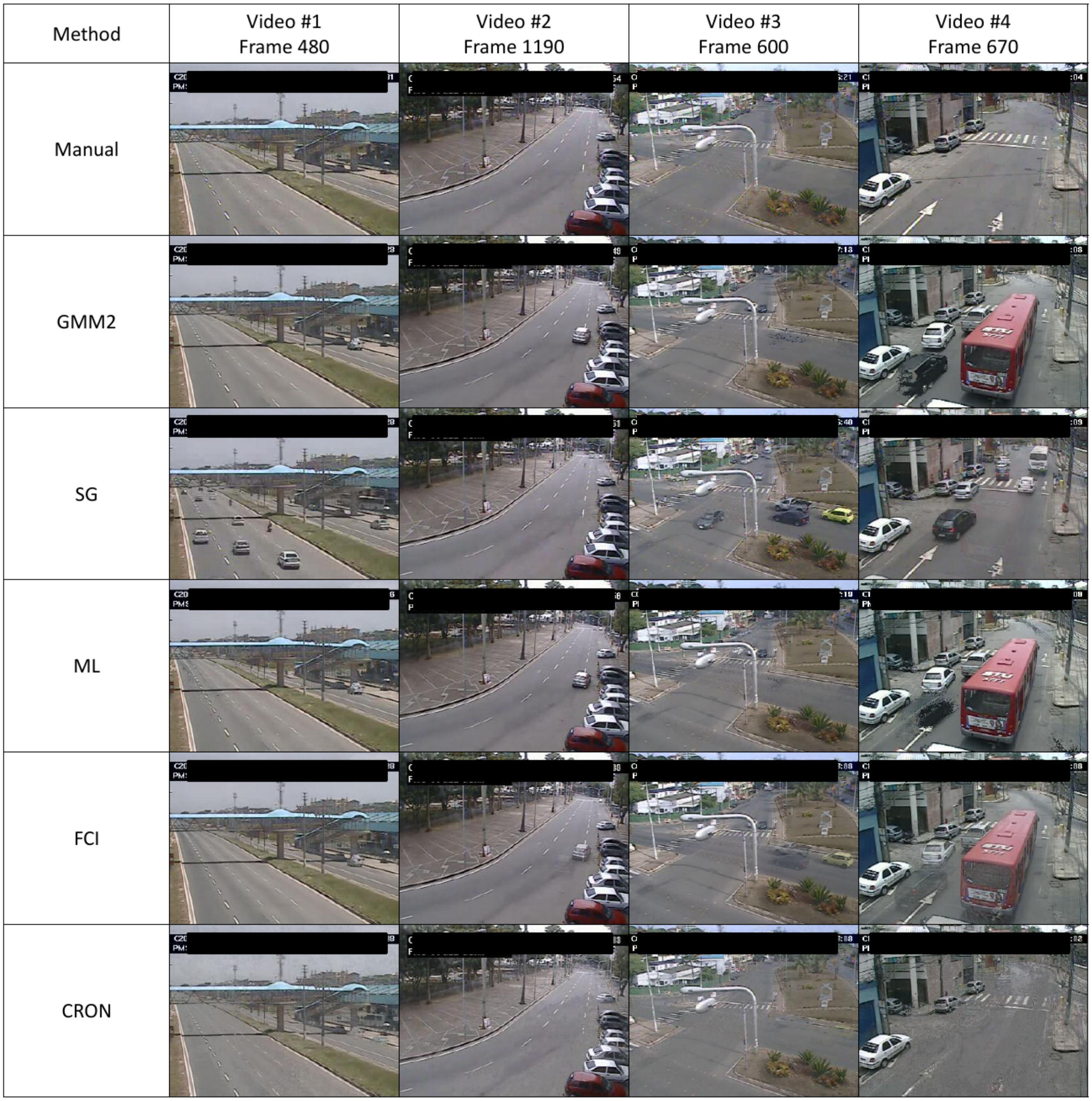


Fig. 7. Background models. In the first row, there are manually extracted background references for each video in a specific frame; the other rows illustrates some examples of the background models given by the best four compared methods and CRON.