VIDEO SURVEILLANCE FOR ROAD TRAFFIC MONITORING

Gonzalo Benito, M. Cristina Bustos, Xián López and Laura Pérez

Computer Vision Center
Universitat Autònoma de Barcelona
{*gonzalo.benito, cristina.bustos, xian.lopez, laura.perezmay*}@*e-campus.uab.cat*

ABSTRACT

Current exponential growth of traffic congestion has proven a major challenge in terms of safety, both in cities and roads alike. Traffic accidents is one of the main death cause worldwide, many efforts are being made in order to reduce it. From a preventive standpoint, traffic control becomes essential not only to make streets and roads safer, but also to make transportation more fluid. In this work, we present a robust and real-time road traffic monitoring system using a regular video camera mounted on a relatively high place. The work is divided in four stages: video stabilization, background modeling and foreground segmentation, vehicle tracking and speed estimation. The system deals with lighting changes and multimodal distributions caused by elements like tree branches or shadows, and it is able to quickly recover from occlusions and track loss. It has been proven to track vehicles successfully in three different real traffic sequences, with distinct cameras and lightning conditions.

Index Terms— road monitoring, background modeling, foreground segmentation, video stabilization, Kalman filter, vehicle tracking, speed estimator

1. MOTIVATION

By 2025, it is predicted that 6.2 billion private motorized trips will be made all around the world every day [1]. This means more roads, more users and many more security concerns. Road traffic monitoring is not just a huge challenge, but a necessity, and computer vision is playing an important role in addressing this problem. In this work we present a real-time road traffic monitoring system using a regular video camera mounted on a relatively high place, i.e. a bridge, with a significant image analysis field. The system is able to track and count vehicles, raising an alarm whenever any of them exceeds the speed limit.

The remaining of this paper is organized as follows. In section 2 different road traffic monitoring methods are presented. In section 3 we present the strategy followed to address the challenge of road traffic surveillance studied along this research. In section 4 the results of our approach are presented and discussed. Finally, in section 5 the conclusions are drawn along with the directions for future research.

2. RELATED WORK

In literature, many approaches have been studied to address the problem of vehicle tracking and speed estimation. On the one hand, we find multimodal approaches. Just to cite a few works published this year, in [1] and [2] the authors use two categories of sensors including IR Lidar and IR camera, and they fuse three detection techniques —Time of Flight (ToF) based, vision based and Laser spot flow based— in a Kalman filter framework. In [3] the authors propose a system based on tri-axial anisotropic magnetoresistive sensors and wireless sensor network, and in [4] the authors use three AMR sensors that give height of vehicle and speed estimation and number of vehicle passing near range of sensors scheme.

On the other hand, many approaches are based solely on video cameras which benefit greatly from the recent advances in automated video analysis. Just to name a few, in [5], the authors perform region based detection and featured based tracking. Similarly, in [6] the authors detect moving vehicles through background/foreground segmentation techniques and estimate vehicles speed per class using feature tracking and nearest neighbors algorithms. In [7], vehicle detection is based on motion detection, using kernel density estimation in a pixel-based technique. Furthermore, the authors use a 3D pose estimation to handle occlusions, and a Kalman filter to perform the tracking. [8] presents a pyramidal approach of Lucas-Kanade algorithm to estimate motion vectors, and tracking of moving objects is performed using Kalman filter, tracking single or multiple moving objects.

3. METHODOLOGY

Our main goal is to create a robust video surveillance road traffic monitoring system able to track vehicles, count them and estimate their speed, raising an alarm if any of them exceeds the speed limit established for the road.

[1] `https://goo.gl/7YcmFW`

Our system is based entirely on motion, and its complete pipeline is described in Fig. 1. The first step, in case the sequence is jittery, is to perform video stabilization using Target Tracking. Then, the foreground is segmented using Stauffer & Grimson algorithm [9], and morphological operations are applied to improve the segmentation. Once a series of blobs have been detected on each frame, a Kalman Filter is used to track them along the video sequence. Finally, using some knowledge of the scene geometry, the vehicles are counted and their speeds are estimated, raising an alarm if any of them exceeds the —known— limit.

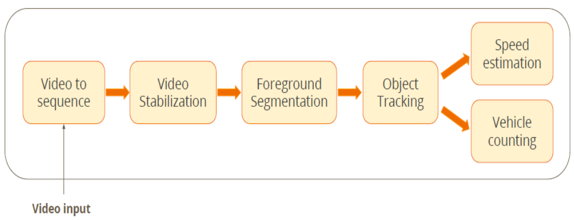

Fig. 1. Video Surveillance for Road Traffic Monitoring System pipeline.

The whole system has been implemented in Matlab, and the code is available at our Github repository [2].

3.1. Target Tracking Video Stabilization

Video stabilization is a family of techniques used to reduce blurring associated with the motion of the camera during exposure, and it generally compensates for pan and tilt of the camera, although rotations can also be compensated.

The main idea of Target Tracking Video Stabilization [3] is to use a block-based parametric motion model to correct translational and rotational camera motions. We define the target to track (in our system a square area in the lower left corner) and we establish a dynamic search region, whose position is determined by the last known target location. Then, we search for the target only within this search region. In each subsequent video frame we determine how much the target has moved relative to the previous frame, and we use this information to remove unwanted translational camera motions and generate a stabilized video. Notice that we add a black padding to all the frames to deal with the loss of information in the borders after stabilization, and we apply the same stabilization and padding to the ground truth for evaluation.

[2]https://github.com/mcv-m4-video/mcv-m4-2017-team2

[3]https://es.mathworks.com/help/vision/examples/video-stabilization.html

Fig. 2. Target Tracking Video Stabilization frame. The square on the left bottom area denotes the area to track used by the stabilization algorithm.

3.2. Background Subtraction & Foreground Detection

Both background and foreground are modeled using Gaussian Mixture Models (GMM), as proposed by Stauffer and Grimson [9], a reliable technique that offers robustness against lighting changes, repetitive motions and long-term changes in the scene. It is based on the temporal observation of the pixels, and several Gaussians are used to represent either foreground or background. Gaussians are considered as background if they have high weight and low variance, and as foreground if they have low weight and high variance, and a pixel is assigned as background or foreground depending on its nearest Gaussian.

Once a foreground segmentation has been obtained, morphological operators are applied to improve the binary mask. First, closing and opening operators are applied to get better defined blobs. Then, a filling holes algorithm is used to avoid holes in the blobs. And last, an area filtering is used to remove the possible noise. Fig. 3 shows an example of foreground detection for a frame of the highway sequence.

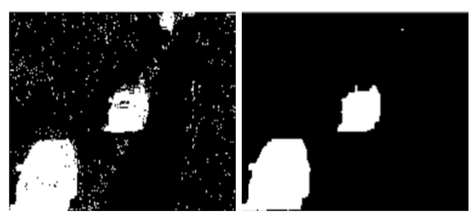

Fig. 3. Foreground segmentation. On the left, the mask obtained with S&G. On the right, the same mask after applying morphological operators.

3.3. Tracking Vehicles

In order to achieve the goals of the system, vehicle counting and speed estimation, it is necessary to keep track of the vehicles in the scene. This means knowing how the detected blobs relate to each other between consecutive frames.

For this purpose, we assume a Linear Dynamics Model (LDM) and, consequently, a Kalman filter [10] is used for tracking. Under this framework, the state of each vehicle is supposed to vary linearly on each time step and suffer a slight perturbation (with Gaussian distribution). Similarly, the detection obtained is supposed to be a linear transformation of the previous state, plus some Gaussian noise as well. The procedure consists of making a prediction of the object's next state, based on all the previous measurements, and once the new detection is done, update this prediction. The equations derived by the Kalman filter are the optimal ones in the case of an LDM.

The assumption of a linear model is not so restrictive as it may seem. This can handle, for instance, the projection (into the camera plane) of the movement of cars traveling at a constant speed. In almost any road it is possible to find sectors with this characteristics. A 50 meter long straight part of the road would be enough for our system to work. Fig. 4 shows an example of kalman tracking for 2 vehicle in our sequence.

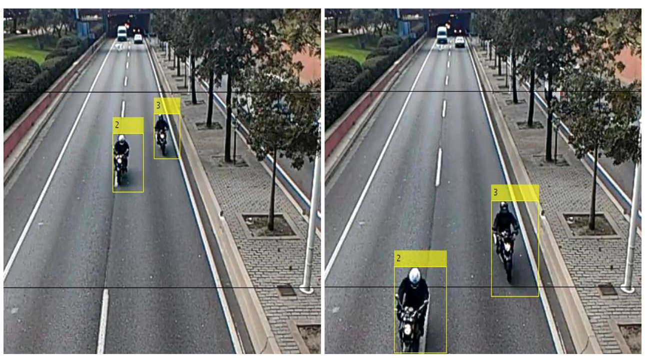

Fig. 4. Sample of Kalman tracking detection for two vehicles along different frames.

3.4. Speed Estimator

Our system provides an estimation on the speed of each tracked vehicle, rising a visual alarm when the speed limits are exceeded. Other works on this field make use of the optical flow in order to make motion estimations, but this approach requires the conversion of pixels/frame to distance/time. Even though it is a feasible task, prior knowledge of pixels/distance ratio in the sequence is mandatory. Moreover, perspective generates a progressive distortion in this ratio along the regions of the image, which can be mitigated by the use of homography techniques. The need for a controlled test with a vehicle for calibration stands as a major drawback.

To avoid the above-mentioned problems we came up with a more straightforward approach based on a single estimation per track, reference points in the view field, and a known distance. This method does not need much calibration to achieve good results, and it requires fewer resources making its implementation simpler. We set 2 reference lines as seen in Fig. 5 in the road and we compute the frame difference between the vehicle crossing the first marker and reaching the second one. Knowing the frame rate of the video and the distance between markers, speed is calculated by a simple equation.

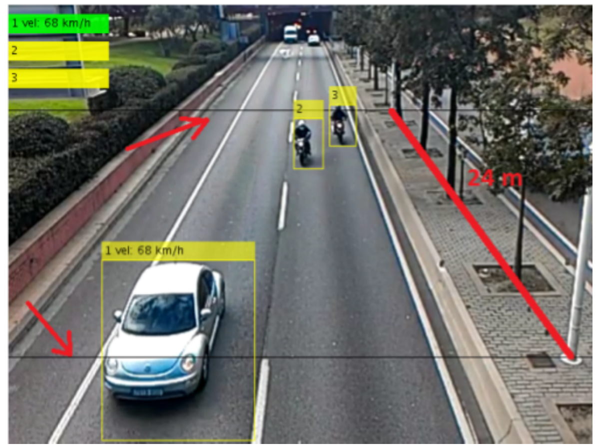

Fig. 5. System speed estimation layout with markers and distance pointed in red.

This method relies on the assumption that the tracked detection is precise and the bounding box is considered the vehicle itself. This avoids the use of the detection's centroid, which introduces greater distortion in the relative distance to the marker line than a border does. It is most desirable that the vanishing point lies out of frame for the perspective to be less distorting. This assumption approximates the real case scenario quite well provided that the distance between markers is as big as possible. The further away the markers are with respect to each other, the less influence a distance error will have on the estimated speed.

4. EVALUATION

The development of our system implied separate tests over each subsystem: stabilization, foreground detection, tracking and speed estimation. This means that we had to perform parameters fine tunning and for this purpose we made use of two dummy video sequences from CDNET [11] that have ground truth attached. As seen in Fig. 6 both sequences *Traffic* and *Highway* consist of road takes, one of them being severely affected by jitter so as to provide a better stabilization test with the method explained in 3.1.

For the final adjustments and field tests, we made a video of our own on a specific road, seen in the last frame of Fig. 6.

Sequence	# Gaussians	# Training Frames	Learning rate	Minimum Background ratio
Traffic	2	10	0.025	0.8
Highway	3	100	0.0025	0.6
Our sequence	2	25	0.0025	0.9

Table 1. Stauffer and Grimson parameters for the three tested video sequences.

Fig. 6. Samples of the tested Sequences. Left: Traffic. Center: Highway, Right: Icària sequence (our sequence)

The chosen location was *Parc de la Nova Icària* within the city of Barcelona, following two main reasons. First, the area is a limited speed sector so it makes easier to evaluate our speed estimation performance. Second, the place has bridges over the road which allows us to locate the camera with an advantageous viewpoint for the system's operation.

Table 1 displays the configurations on the Stauffer & Grimson algorithm with the best performance on our tests. Said performance was assessed making use of the ground truth provided by CDNET, analyzing precision, recall and F1 index which results can be seen in Table 2.

Our speed estimation also shows coherent results. In the case of *Traffic* sequence, results are exposed in Table 3, only the first vehicle's speed is rather low compared to the expected 80 km/h for a road. The vehicles from *Highway* sequence were also estimated to have speeds in accordance with expectations, estimated speeds can be seen in Table 4. In the case of our own video sequence, the higher number of vehicles passing by make it more convenient to show results with the histogram in Fig. 7. As it can be seen, speeds lie mostly between 60 and 75 km/h, something usual for an 80 km/h maximum speed sector coming out of a tunnel. Overall, the test results are positive and demonstrate a good performance of the system in general.

Sequence	Precision	Recall	F1
Traffic	0.65	0.71	0.65
Highway	0.90	0.84	0.86

Table 2. Stauffer & Grimson foreground segmentation results for the Traffic and Highway sequences.

Vehicle	Speed (km/h)
1	50
2	84
3	76

Table 3. Speed estimation. Estimated speed for the vehicles tracked on the *Traffic* sequence

Vehicle	Speed (km/h)
1	-
2	89
3	89
4	86
5	81
6	81

Table 4. Speed estimation. Estimated speed for the vehicles tracked on the *Highway* sequence

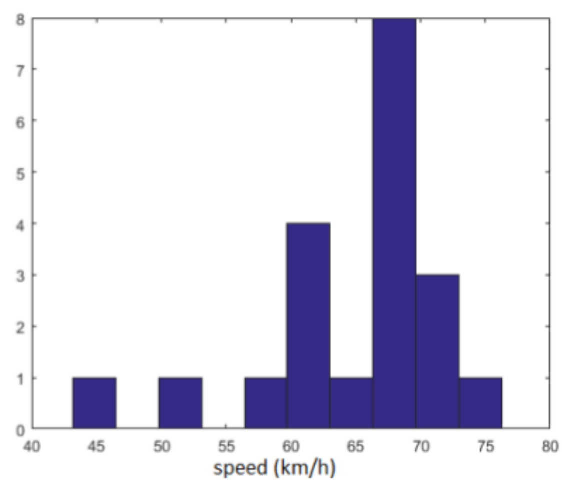

Fig. 7. Speed estimation. Histogram of the registered speeds for Icària video sequence.

5. CONCLUSIONS

This paper has shown a complete road traffic monitoring system in four stages: video stabilization, background modeling and foreground segmentation, vehicle tracking and speed estimation. We showed a robust and stable method able to work in real-time, which requires only the Stauffer and Grimson parameters —robust to different cameras and different scenes—, a known distance and the camera frame rate for speed estimation and the speed limit to raise the appropriate alarm whenever any vehicle exceeds it. As it is based in Stauffer and Grimson [9], the system deals with slow lighting changes by progressively adapting the values of the Gaussians, and it also deals with multi-modal distributions caused by shadows, swaying branches and other features of the real world which are often troublesome. Furthermore, it is able to quickly recover from occlusions and track loss.

The system has been successfully used to track vehicles in three different real traffic sequences in outdoor environments, involving different cameras and different lighting conditions, raising and alarm whenever any vehicle exceed the speed limit, and it achieves our goals of real time road traffic monitoring.

6. REFERENCES

[1] Yi Tian, Hangxin Liu, and Tomonari Furukawa, "Reliable infrastructural urban traffic monitoring via lidar and camera fusion," *SAE International Journal of Passenger Cars-Electronic and Electrical Systems*, vol. 10, no. 2017-01-0083, 2017.

[2] Yi Tian, *Self-Powered Intelligent Traffic Monitoring Using IR Lidar and Camera*, Ph.D. thesis, Virginia Tech, 2017.

[3] Qifan Wei et al., "Adaptable vehicle detection and speed estimation for changeable urban traffic with anisotropic magnetoresistive sensor," *IEEE Sensors Journal*, 2017.

[4] Surabhi P Jinturkar and Sushant J Pawar, "Real-time traffic monitoring with portable amr sensor system," in *International Conference on Smart Trends for Information Technology and Computer Communications*. Springer, 2016, pp. 408–417.

[5] Mohamed H Zaki, Tarek Sayed, and Shewkar E Ibrahim, "Demonstrating the viability of automated video analysis for mid-block traffic data collection: A study in the city of edmonton," Tech. Rep., 2017.

[6] Agustín Yabo, Sebastián I Arroyo, Félix G Safar, and Damián Oliva, "Vehicle classification and speed estimation using computer vision techniques," in *XXV Congreso Argentino de Control Automático (AADECA 2016)(Buenos Aires, 2016)*, 2016.

[7] Rashmita Khilar and S Chitrakala, "A novel method for effective vehicle detection and tracking eliminating occlusion," *Asian Journal of Research in Social Sciences and Humanities*, vol. 6, no. cs1, pp. 714–727, 2016.

[8] Bhagyashri B More and Manasi Pathade, "Moving object detection, tracking, object counting and speed measurement," .

[9] Chris Stauffer and W Eric L Grimson, "Adaptive background mixture models for real-time tracking," in *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on.* IEEE, 1999, vol. 2, pp. 246–252.

[10] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Transactions of the ASME - Journal of Basic Engineering*, vol. 82, pp. 25–45, 1960.

[11] N. Goyette, P.-M. Jodoin, F. Porikli, J. Konrad, and P. Ishwar, "changedetection.net: A new change detection benchmark dataset," in *IEEE Workshop on Change Detection (CDW-2012) at CVPR-2012, Providence, RI*, 2012.